## (19) World Intellectual Property Organization International Bureau



### 

## (43) International Publication Date 21 March 2002 (21.03.2002)

### PCT

## (10) International Publication Number WO 02/22080 A2

(51) International Patent Classification7:

A61K

- (21) International Application Number: PCT/US01/28861
- (22) International Filing Date:

14 September 2001 (14.09.2001)

(25) Filing Language:

English

(26) Publication Language:

English

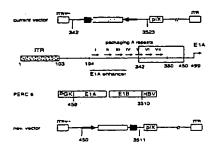
- (30) Priority Data: 60/233,180 15 September 2000 (15.09.2000) U
- (71) Applicant (for all designated States except US): MERCK & CO., INC. [US/US]; 126 East Lincoln Avenue, Rahway, NJ 07065-0907 (US).
- (72) Inventors; and
- (75) Inventors/Applicants (for US only): EMINI, Emilio, A. [US/US]; 126 East Lincoln Avenue, Rahway, NJ 07065-0907 (US). YOUIL, Rima [AU/US]; 126 East Lincoln Avenue, Rahway, NJ 07065-0907 (US). BETT, Andrew, J.

[CA/US]; 126 East Lincoln Avenue, Rahway, NJ 07065-0907 (US). CHEN, Ling [US/US]; 126 East Lincoln Avenue, Rahway, NJ 07065-0907 (US). KASLOW, David, C. [US/US]; 126 East Lincoln Avenue, Rahway, NJ 07065-0907 (US). SHIVER, John, W. [US/US]; 126 East Lincoln Avenue, Rahway, NJ 07065-0907 (US). TONER, Timothy, J. [US/US]; 126 East Lincoln Avenue, Rahway, NJ 07065-0907 (US). CASIMIRO, Daniel, R. [PH/US]; 126 East Lincoln Avenue, Rahway, NJ 07065-0907 (US).

- (74) Common Representative: MERCK & CO., INC.; 126 East Lincoln Avenue, Rahway, NJ 07065-0907 (US).
- (81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.
- (84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian

[Continued on next page]

(54) Title: ENHANCED FIRST GENERATION ADENOVIRUS VACCINES EXPRESSING CODON OPTIMIZED HIV1-GAG, POL, NEF AND MODIFICATIONS



Modifications made to the current adenovector backbone in the generation of the ner

(57) Abstract: First generation adenoviral vectors and associated recombinant adenovirus-based HIV vaccines which show enhanced stability and growth properties and greater cellular-mediated immunity are described within this specification. These adenoviral vectors are utilized to generate and produce through cell culture various adenoviral-based HIV-1 vaccines which contain HIV-1 gag, HIV-1 pol and/or HIV-1 nef polynucleotide pharmaceutical products; and biologically relevant modifications thereof. These adenovirus vaccines, when directly introduced into living vertebrate tissue, preferably a mammalian host such as a human or a non-human mammal of commercial or domestic veterinary importance, express the HIV1-Gag, Pol and/or Nef protein or biologically modification thereof, inducing a cellular immune response which specifically recognizes HIV-1. The exemplified polynucleotides of the present invention are synthetic DNA molecules encoding HIV-1 Gag, encoding codon optimized HIV-1 Pol, derivatives of optimized HIV-1 Pol (including constructs wherein protease, reverse transcriptase, RNAse H and integrase activity of HIV-1 Pol is inactivated), HIV-1 Nef and derivatives of optimized HIV-1 Nef, including nef mutants which effect wild type characteristics of Nef, such as myristylation and down regulation of host CD4. The adenoviral vaccines of the present invention, when administered alone or in a combined modality regime, will offer a prophylactic advantage to previously uninfected individuals and/or provide a therapeutic effect by reducing viral load levels within an infected individual, thus prolonging the asymptomatic phase of HIV-1 infection.



patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

#### Published:

 without international search report and to be republished upon receipt of that report

### TITLE OF THE INVENTION

ENHANCED FIRST GENERATION ADENOVIRUS VACCINES EXPRESSING CODON OPTIMIZED HIV1-GAG, POL, NEF AND MODIFICATIONS

### 5 CROSS-REFERENCE TO RELATED APPLICATIONS

This application claims the benefit, under 35 U.S.C. §119(e), of U.S. provisional applications 60/233,180, 60/279,056, and Attorney Docket 20867PV2 (serial number unassigned), filed September 15, 2000, March 27, 2001, and September 7, 2001, respectively.

10

20

25

30

35

# STATEMENT REGARDING FEDERALLY-SPONSORED R&D Not Applicable

### REFERENCE TO MICROFICHE APPENDIX

15 Not Applicable

### FIELD OF THE INVENTION

The present invention relates to recombinant, replication-deficient first generation adenovirus vaccines found to exhibit enhanced growth properties and greater cellular-mediated immunity as compared to other replication-deficient vectors. The invention also relates to the associated first generation adenoviral vectors described herein, which, through the incorporation of additional 5' adenovirus sequence, enhance large scale production efficiency of the recombinant, replicationdefective adenovirus described herein. Another aspect of the instant-invention is the surprising discovery that the intron A portion of the human cytomegalovirus (hCMV) promoter constitutes a region of instability in adenoviral vector constructs. Removal of this region from adenoviral expression constructs results in greatly improved vector stability. Therefore, improved vectors expressing a transgene under the control of an intron A-deleted CMV promoter constitute a further aspect of this invention. These adenoviral vectors are useful for generating recombinant adenovirus vaccines against human immunodeficiency virus (HTV). In particular, the first generation adenovirus vectors disclosed herein are utilized to construct and generate adenovirus-based HIV-1 vaccines which contain HIV-1 Gag, HIV-1 Pol and/or HIV-1 Nef polynucleotide pharmaceutical products, and biologically active modifications thereof. Host administration of the recombinant, replication-deficient adenovirus vaccines described herein results in expression of HIV-1 Gag, HIV-1- Pol and/or Nef protein or

immunologically relevant modifications thereof, inducing a cellular immune response which specifically recognizes HIV-1. The exemplified polynucleotides of the present invention are synthetic DNA molecules encoding codon optimized HIV-1 Gag, HIV-1 Pol, derivatives of optimized HIV-1 Pol (including constructs wherein protease, reverse transcriptase, RNAse H and integrase activity of HIV-1 Pol is inactivated), HIV-1 Nef, and derivatives of optimized HIV-1 Nef, including nef mutants which effect wild type characteristics of Nef, such as myristylation and down regulation of host CD4. The HIV adenovirus vaccines of the present invention, when administered alone or in a combined modality and/or prime/boost regimen, will offer a prophylactic advantage to previously uninfected individuals and/or provide a therapeutic effect by reducing viral load levels within an infected individual, thus prolonging the asymptomatic phase of HIV-1 infection.

### **BACKGROUND OF THE INVENTION**

5

10

15

20

25

30

Human Immunodeficiency Virus-1 (HIV-1) is the etiological agent of acquired human immune deficiency syndrome (AIDS) and related disorders. HIV-1 is an RNA virus of the Retroviridae family and exhibits the 5'LTR-gag-pol-env-LTR 3' organization of all retroviruses. The integrated form of HIV-1, known as the provirus, is approximately 9.8 Kb in length. Each end of the viral genome contains flanking sequences known as long terminal repeats (LTRs). The HIV genes encode at least nine proteins and are divided into three classes; the major structural proteins (Gag, Pol, and Env), the regulatory proteins (Tat and Rev); and the accessory proteins (Vpu, Vpr, Vif and Nef).

The gag gene encodes a 55-kilodalton (kDa) precursor protein (p55) which is expressed from the unspliced viral mRNA and is proteolytically processed by the HIV protease, a product of the pol gene. The mature p55 protein products are p17 (matrix), p24 (capsid), p9 (nucleocapsid) and p6.

The pol gene encodes proteins necessary for virus replication; a reverse transcriptase, a protease, integrase and RNAse H. These viral proteins are expressed as a Gag-Pol fusion protein, a 160 kDa precursor protein which is generated via a ribosomal frame shifting. The viral encoded protease proteolytically cleaves the Pol polypeptide away from the Gag-Pol fusion and further cleaves the Pol polypeptide to the mature proteins which provide protease (Pro, P10), reverse transcriptase (RT, P50), integrase (IN, p31) and RNAse H (RNAse, p15) activities.

The *nef* gene encodes an early accessory HIV protein (Nef) which has been shown to possess several activities such as down regulating CD4 expression, disturbing T-cell activation and stimulating HIV infectivity.

The *env* gene encodes the viral envelope glycoprotein that is translated as a 160-kilodalton (kDa) precursor (gp160) and then cleaved by a cellular protease to yield the external 120-kDa envelope glycoprotein (gp120) and the transmembrane 41-kDa envelope glycoprotein (gp41). Gp120 and gp41 remain associated and are displayed on the viral particles and the surface of HIV-infected cells.

5

10

15

20

25

30

35

The *tat* gene encodes a long form and a short form of the Tat protein, a RNA binding protein which is a transcriptional transactivator essential for HIV-1 replication.

The *rev* gene encodes the 13 kDa Rev protein, a RNA binding protein. The Rev protein binds to a region of the viral RNA termed the Rev response element (RRE). The Rev protein promotes transfer of unspliced viral RNA from the nucleus to the cytoplasm. The Rev protein is required for HIV late gene expression and in turn, HIV replication.

Gp120 binds to the CD4/chemokine receptor present on the surface of helper T-lymphocytes, macrophages and other target cells in addition to other co-receptor molecules. X4 (macrophage tropic) virus show tropism for CD4/CXCR4 complexes while a R5 (T-cell line tropic) virus interacts with a CD4/CCR5 receptor complex. After gp120 binds to CD4, gp41 mediates the fusion event responsible for virus entry. The virus fuses with and enters the target cell, followed by reverse transcription of its single stranded RNA genome into the double-stranded DNA via a RNA dependent DNA polymerase. The viral DNA, known as provirus, enters the cell nucleus, where the viral DNA directs the production of new viral RNA within the nucleus, expression of early and late HIV viral proteins, and subsequently the production and cellular release of new virus particles. Recent advances in the ability to detect viral load within the host shows that the primary infection results in an extremely high generation and tissue distribution of the virus, followed by a steady state level of virus (albeit through a continual viral production and turnover during this phase), leading ultimately to another burst of virus load which leads to the onset of clinical AIDS. Productively infected cells have a half life of several days, whereas chronically or latently infected cells have a 3-week half life, followed by non-productively infected cells which have a long half life (over 100 days) but do not significantly contribute to day to day viral loads seen throughout the course of disease.

Destruction of CD4 helper T lymphocytes, which are critical to immune defense, is a major cause of the progressive immune dysfunction that is the hallmark of HIV infection. The loss of CD4 T-cells seriously impairs the body's ability to fight most invaders, but it has a particularly severe impact on the defenses against viruses, fungi, parasites and certain bacteria, including mycobacteria.

5

10

15

20

25

30

35

Effective treatment regimens for HIV-1 infected individuals have become available recently. However, these drugs will not have a significant impact on the disease in many parts of the world and they will have a minimal impact in halting the spread of infection within the human population. As is true of many other infectious diseases, a significant epidemiologic impact on the spread of HIV-1 infection will only occur subsequent to the development and introduction of an effective vaccine. There are a number of factors that have contributed to the lack of successful vaccine development to date. As noted above, it is now apparent that in a chronically infected person there exists constant virus production in spite of the presence of anti-HIV-1 humoral and cellular immune responses and destruction of virally infected cells. As in the case of other infectious diseases, the outcome of disease is the result of a balance between the kinetics and the magnitude of the immune response and the pathogen replicative rate and accessibility to the immune response. Pre-existing immunity may be more successful with an acute infection than an evolving immune response can be with an established infection. A second factor is the considerable genetic variability of the virus. Although anti-HIV-1 antibodies exist that can neutralize HIV-1 infectivity in cell culture, these antibodies are generally virus isolate-specific in their activity. It has proven impossible to define serological groupings of HIV-1 using traditional methods. Rather, the virus seems to define a serological "continuum" so that individual neutralizing antibody responses, at best, are effective against only a handful of viral variants. Given this latter observation, it would be useful to identify immunogens and related delivery technologies that are likely to elicit anti-HIV-1 cellular immune responses. It is known that in order to generate CTL responses antigen must be synthesized within or introduced into cells, subsequently processed into small peptides by the proteasome complex, and translocated into the endoplasmic reticulum/Golgi complex secretory pathway for eventual association with major histocompatibility complex (MHC) class I proteins. CD8<sup>+</sup> T lymphocytes recognize antigen in association with class I MHC via the T cell receptor (TCR) and the CD8 cell surface protein. Activation of naive CD8<sup>+</sup> T cells into activated effector or memory cells generally requires both TCR engagement of antigen as described above as well as engagement of costimulatory proteins. Optimal

induction of CTL responses usually requires "help" in the form of cytokines from CD4<sup>+</sup> T lymphocytes which recognize antigen associated with MHC class II molecules via TCR and CD4 engagement.

5

10

15

20

25

30 .

35

European Patent Applications 0 638 316 (Published February 15, 1995) and 0 586 076 (Published March 9, 1994), (both assigned to American Home Products Corporation) describe replicating adenovirus vectors carrying an HIV gene, including *env* or *gag*. Various treatment regimens were used with chimpanzees and dogs, some of which included booster adenovirus or protein plus alum treatments.

Replication-defective adenoviral vectors harboring deletions in the E1 region are known, and recent adenoviral vectors have incorporated the known packaging repeats into these vectors; e.g., see EP 0 707 071, disclosing, *inter alia*, an adenoviral vector deleted of E1 sequences from base pairs 459 to 3328; and U.S. Patent No. 6,033,908, disclosing, *inter alia*, an adenoviral vector deleted of base pairs 459-3510. The packaging efficiency of adenovirus has been taught to depend on the number of incorporated individual A (packaging) repeats; *see*, *e.g.*, Gräble and Hearing, 1990 *J. Virol*. 64(5):2047-2056; Gräble and Hearing, 1992 *J. Virol*. 66(2):723-731.

Larder, et al., (1987, *Nature* 327: 716-717) and Larder, et al., (1989, *Proc. Natl. Acad. Sci.* 86: 4803-4807) disclose site specific mutagenesis of HIV-1 RT and the effect such changes have on *in vitro* activity and infectivity related to interaction with known inhibitors of RT.

Davies, et al. (1991, *Science* 252:, 88-95) disclose the crystal structure of the RNase H domain of HIV-1 Pol.

Schatz, et al. (1989, FEBS Lett. 257: 311-314) disclose that mutations Glu478Gln and His539Phe in a complete HIV-1 RT/RNase H DNA fragment results in defective RNase activity without effecting RT activity.

Mizrahi, et al. (1990, *Nucl. Acids. Res.* 18: pp. 5359-5353) disclose additional mutations Asp443Asn and Asp498Asn in the RNase region of the *pol* gene which also results in defective RNase activity. The authors note that the Asp498Asn mutant was difficult to characterize due to instability of this mutant protein.

Leavitt, et al. (1993, *J. Biol. Chem.* 268: 2113-2119) disclose several mutations, including a Asp64Val mutation, which show differing effect on HIV-1 integrase (IN) activity.

Wiskerchen, et al. (1995, *J. Virol*. 69: 376-386) disclose singe and double mutants, including mutation of aspartic acid residues which effect HIV-1 IN and viral replication functions.

It would be of great import in the battle against AIDS to produce a prophylactic- and/or therapeutic-based HIV vaccine which generates a strong cellular immune response against an HIV infection. The present invention addresses and meets these needs by disclosing a class of adenovirus vaccines which, upon host administration, express codon optimized and modified versions of the HIV-1 genes, gag, pol and nef. These recombinant, replication-defective adenovirus vaccines may be administered to a host, such as a human, alone or as part of a combined modality regimen and/or prime-boost vaccination regimen with components of the present invention and/or a distinct viral HIV DNA vaccine, non-viral HIV DNA vaccine, HIV subunit vaccine, an HIV whole killed vaccine and/or a live attenuated HIV vaccine.

### SUMMARY OF THE INVENTION

5

10

15

20

25

30

35

The present invention relates to enhanced replication-defective recombinant adenovirus vaccine vectors and associated recombinant, replication-deficient adenovirus vaccines which encode various forms of HIV-1 Gag, HIV-1 Pol, and/or HIV-1 Nef, including immunologically relevant modifications of HIV-1 Gag, HIV-1 Pol and HIV-1 Nef. The adenovirus vaccines of the present invention express HIV antigens and provide for improved cellular-mediated immune responses upon host administration. Potential vaccinees include but are not limited to primates and especially humans and non-human primates, and also include any non-human mammal of commercial or domestic veterinary importance. An effect of the improved recombinant adenovirus-based vaccines of the present invention should be a lower transmission rate to previously uninfected individuals (i.e., prophylactic applications) and/or reduction in the levels of the viral loads within an infected individual (i.e., therapeutic applications), so as to prolong the asymptomatic phase of HIV-1 infection. In particular, the present invention relates to adenoviral-based vaccines which encode various forms of codon optimized HIV-1 Gag (including but in no way limited to p55 versions of codon optimized full length (FL) Gag and tPA-Gag fusion proteins), HIV-1 Pol, HIV-1 Nef, and selected modifications of immunological relevance. The administration, intracellular delivery and expression of these adenovirus vaccines elicit a host CTL and Th response. The preferred replication-defective recombinant adenoviral vaccine vectors include but are not limited to synthetic DNA molecules which (1) encode codon optimized versions of wild type HIV-1 Gag; (2) encode codon optimized versions of HIV-1 Pol; (3) encode codon optimized versions of HIV-1 Pol fusion proteins; (4) encode codon optimized versions of modified HIV-1 Pol proteins and fusion proteins, including but not limited

to pol modifications involving residues within the catalytic regions responsible for RT, RNase and IN activity within the host cell; (5) encode codon optimized versions of wild type HIV-1 Nef; (6) codon optimized versions of HIV-1 Nef fusion proteins; and/or (7) codon optimized versions of HIV-1 Nef derivatives, including but not limited to nef modifications involving introduction of an amino-terminal leader sequence, removal of an amino-terminal myristylation site and/or introduction of dileucine motif mutations. The Nef-based fusion and modified proteins, disclosed within this specification and expressed from an adenoviral-based vector vaccine this specification, may possess altered trafficking and/or host cell function while retaining the ability to be properly presented to the host MHC I complex and in turn elicit a host CTL and Th response. Examples of HIV-1 Gag, Pol and/or Nef fusion proteins include but are not limited to fusion of a leader or signal peptide at the NH<sub>2</sub>-teriminal portion of the viral antigen coding region. Such a leader peptide includes but is not limited to a tPA leader peptide.

10

15

20

25

30

35

The adenoviral vector utilized in construction of the HIV-1 Gag-, HIV-1 Poland/or HIV-1 Nef- based vaccines of the present invention may comprise any replication-defective adenoviral vector which provides for enhanced genetic stability of the recombinant adenoviral genome through large scale production and purification of the recombinant virus. In other words, an HIV-1 Gag-, Pol- or Nef-based adenovirus vaccine of the present invention is a purified recombinant, replicationdefective adenovirus which is shown to be genetically stable through multiple passages in cell culture and remains so during large scale production and purification procedures. Such a recombinant adenovirus vector and harvested adenovirus vaccine lends itself to large scale dose filling and subsequent worldwide distribution procedures which will be demanded of an efficacious monovalent or multivalent HIV vaccine. The present invention meets this basic requirement with description of a replication-defective adenoviral vector and vectors derived therefrom, at least partially deleted in E1, comprising a wildtype adenovirus cis-acting packaging region from about base pair 1 to between from about base pair 342 (more preferably, 400) to about base pair 458 of the wildtype adenovirus genome. A preferred embodiment of the instant invention comprises base pairs 1-450 of a wildtype adenovirus. In other preferred embodiments, the replication -defective adenoviral vector has, in addition thereto, a region 3' to the E1-deleted region comprising base pairs 3511-3523. Basepairs 342-450 (more particularly, 400-450) constitute an extension of the 5'region of previously disclosed vectors carrying viral antigens, particularly HIV antigens (see, e.g., PCT International Application PCT/US00/18332, published

January 11, 2001 (WO 01/02067), which claims priority to U.S. Provisional Application Serial Nos. 60/142,631 and 60/148,981, filed 7/6/1999 and 8/13/1999, respectively; these documents herein incorporated by reference. Applicants have found that extending the 5' region further into the E1 gene into the disclosed vaccine vectors incorporated elements found to be important in optimizing the packaging of the virus.

As compared to previous vectors not comprising basepairs from about 1 to between from about base pair 342 (more preferably, 400) to about base pair 458 of the wildtype adenovirus genome, vectors comprising the above region exhibited enhanced growth characteristics, with approximately 5-10 fold greater amplification rates, a more potent virus effect, allowing lower doses of virus to be used to generate equivalent immunity; and a greater cellular-mediated immune response than replication-deficient vectors not comprising this region (basepairs 1-450). Even more important, adenoviral constructs derived therefrom are very stable genetically in large-scale production, particularly those comprising an expression cassette under the control of a hCMV promoter devoid of intron A. This is because Applicants have surprisingly found that the intron A portion of the hCMV promoter constituted a region of instability when employed in adenoviral vectors. Applicants have, therefore, identified an enhanced adenoviral vector which is particularly suited for use in gene therapy and nucleotide-based vaccine-vectors which, favorably, lends itself to large scale propagation.

10

15

20

25

30

35

A preferred embodiment of this invention is a replication-defective adenoviral vector in accordance with the above description wherein the gene is inserted in the form of a gene expression cassette comprising (a) a nucleic acid encoding a protein or biologically active and/or immunologically relevant portion thereof; (b) a heterologous promoter operatively linked to the nucleic acid of part a); and, (c) a transcription terminator.

In preferred embodiments, the E1 gene, other than that contained within basepairs 1-450 or, alternatively, that contained within base pairs 1-450 and 3511-3523 has been deleted from the adenoviral vector, and the gene expression cassette has replaced the deleted E1 gene. In other preferred embodiments, the replication defective adenovirus genome does not have a functional E3 gene, or the E3 gene has been deleted. Most preferably, the E3 region is present within the adenoviral genome. Further preferred embodiments are wherein the gene expression cassette is in an E1 anti-parallel (transcribed in a 3' to 5' direction relative to the vector backbone)

orientation or, more preferably, an E1 parallel (transcribed in a 5' to 3' direction relative to the vector backbone) orientation.

Further embodiments relate to a shuttle plasmid vector comprising: an adenoviral portion and a plasmid portion, wherein said adenovirus portion comprises: 5 a) a replication defective adenovirus genome, at least partially deleted in E1, comprising a wildtype adenovirus cis-acting packaging region from about base pair 1 to between from about base pair 342 (more preferably, 400) to about base pair 458 (preferably, 1-450) of the wildtype adenovirus genome and, preferably, in addition thereto, basepairs 3511-3523 of a wildtype adenovirus sequence; and b) a gene expression cassette comprising: (a) a nucleic acid encoding a protein or biologically active and/or immunologically relevant portion thereof; (b) a heterologous promoter operatively linked to the nucleic acid of part a); and (c) a transcription terminator and/or a polyadenylation site.

10

15

20

25

30

35

Other aspects of this invention include a host cell comprising said adenoviral vectors and/or said shuttle plasmid vectors; vaccine compositions comprising said vectors; and methods of producing the vectors comprising (a) introducing the adenoviral vector into a host cell which expresses adenoviral E1 protein, and (b) harvesting the resultant adenoviral vectors.

To this end, the present invention particularly relates to harvested recombinant, replication defective virus derived from a host cell, such as but not limited to 293 cells or PER.C6® cells, including but not limited to harvested virus related to any of the MRKAd5 vector backbones, with or without an accompanying transgene, including but not limited to the HIV-1 antigens described herein. An HTV-1 vaccine is represented by any harvested, recombinant adenovirus material which expresses any one or more of the HIV-1 antigens disclosed herein. This harvested material may then be purified, formulated and stored prior to host administration.

Another aspect of this invention is a method of generating a cellular immune response against a protein in an individual comprising administering to the individual an adenovirus vaccine vector comprising:

a) a recombinant, replication defective adenoviral vector, at least partially deleted in E1, comprising a wildtype adenovirus cis-acting adenovirus packaging region from about base pair 1 to between from about base pair 342 (more preferably, 400) to about base pair 458 (preferably, 1-450) and, preferably in addition thereto, base pairs 3511-3523 of a wildtype adenovirus sequence, and,

b) a gene expression cassette comprising:(i) a nucleic acid encoding a protein or biologically active and/or immunologically relevant portion thereof; (ii) a heterologous promoter operatively linked to the nucleic acid of part a); and (iii) a transcription terminator and/or a polyadenylation site.

5

10

15

20

25

30

35

In view of the efficacious nature of the adenoviral and/or DNA plasmid vaccines described herein, the present invention relates to all methodology regarding administration of one or more of these adenoviral and/or DNA plasmid vaccines to provide effective immunoprophylaxis, to prevent establishment of an HIV-1 infection following exposure to this virus, or as a post-HIV infection therapeutic vaccine to mitigate the acute HIV-1 infection so as to result in the establishment of a lower virus load with beneficial long term consequences. As discussed herein, such a treatment regimen may include a monovalent or multivalent composition, various combined modality applications, and/or a prime/boost regimen to as to optimize antigen expression and a concomitant cellular-mediated and/or humoral immune response upon inoculation into a living vertebrate tissue. Therefore, the present invention provides for methods of using the adenoviral and/or DNA plasmid vaccines disclosed herein within the various parameters disclosed herein as well as any additional parameters known in the art, which, upon introduction into mammalian tissue induces intracellular expression of the gag, pol and/or nef-based vaccines.

To this end, the present invention relates in part to methods of generating a cellular immune response in a vaccinee, preferably a human vaccinee, wherein the individual is given more than one administration of adenovirus vaccine vector, and it may be given in a regimen accompanied by the administration of a plasmid vaccine. The plasmid vaccine (also referred to herein as a "DNA plasmid vaccine" or "vaccine plasmid" comprises a nucleic acid encoding a protein or an immunologically relevant portion thereof, a heterologous promoter operably linked to the nucleic acid sequence, and a transcription terminator or a polyadenylation signal (such as bGH or SPA, respectively). There may be a predetermined minimum amount of time separating the administrations. The individual can be given a first dose of plasmid vaccine, and then a second dose of plasmid vaccine. Alternatively, the individual may be given a first dose of adenovirus vaccine, and then a second dose of adenovirus vaccine. In other embodiments, the plasmid vaccine is administered first, followed after a time by administration of the adenovirus vaccine. Conversely, the adenovirus vaccine may be administered first, followed by administration of plasmid vaccine after a time. In these embodiments, an individual may be given multiple doses of the same adenovirus serotype in either viral vector or plasmid form, or the virus may be of

differing serotypes. In the alternative, a viral antigen of interest can be first delivered via a viral vaccine other than an adenovirus-based vaccine, and then followed with the adenoviral vaccine disclosed. Alternative viral vaccines include but are not limited to pox virus and venezuelan equine encephilitis virus.

5

10

15

20

25

30

35

The present invention also relates to multivalent adenovirus vaccine compositions which comprise Gag, Pol and Nef components described herein; see, e.g., Example 29 and Table 25. Such compositions will provide for an enhanced cellular immune response subsequent to host administration, particularly given the genetic diversity of human MHCs and of circulating virus. Examples, but not limitations, include MRKAd5-vector based multivalent vaccine compositions which provide for a divalent (i.e., gag and nef, gag and pol, or pol and nef components) or a trivalent vaccine (i.e., gag, pol and nef components) composition. Such a mutlivalent vaccine may be filled for a single dose or may consist of multiple inoculations of each individually filled component; and may in addition be part of a prime/boost regimen with viral or non-viral vector vaccines as introduced in the previous paragraph. To this end, preferred compositions are MRKAd5 adenovirus used in combination with multiple, distinct HIV antigen classes. Each HIV antigen class is subject to sequence manipulation, thus providing for a multitude of potential vaccine combinations; and such combinations are within the scope of the present invention. The utilization of such combined modalities vaccine formulation and administration increase the probability of eliciting an even more potent cellular immune response when compared to inoculation with a single modality regimen.

The concept of a "combined modality" as disclosed herein also covers the alternative mode of administration whereby multiple HIV-1 viral antigens may be ligated into a proper shuttle plasmid for generation of a pre-adenoviral plasmid comprising multiple open reading frames. For example, a trivalent vector may comprise a gag-pol-nef fusion, in either a E3(-) or E3(+) background, preferably a E3 deleted backbone, or possibly a "2+1" divalent vaccine, such as a gag-pol fusion (i.e., codon optimized p55 gag and inactivated optimized pol; Example 29 and Table 25) within the same MRKAd5 backbone, with each open reading frame being operatively linked to a distinct promoter and transcription termination sequence. Alternatively, the two open reading frames may be operatively linked to a single promoter, with the open reading frames operatively linked by an internal ribosome entry sequence (IRES). Therefore, a multivalent vaccine delivered as a single, or possibly a second harvested recombinant, replication-deficient adenovirus is contemplated as part of the present invention.

Therefore, the adenoviral vaccines and plasmid DNA vaccines of this invention may be administered alone, or may be part of a prime and boost administration regimen. A mixed modality priming and booster inoculation scheme will result in an enhanced immune response, particularly if pre-existing anti-vector immune responses are present. This one aspect of this invention is a method of priming a subject with the plasmid vaccine by administering the plasmid vaccine at least one time, allowing a predetermined length of time to pass, and then boosting by administering the adenoviral vaccine. Multiple primings typically, 1-4, are usually employed, although more may be used. The length of time between priming and boost may typically vary from about four months to a year, but other time frames may be used. In experiments with rhesus monkeys, the animals were primed four times with plasmid vaccines, then were boosted 4 months later with the adenoviral vaccine. Their cellular immune response was notably higher than that of animals which had only received adenoviral vaccine. The use of a priming regimen may be particularly preferred in situations where a person has a pre-existing anti-adenovirus immune response.

5

10

15

20

25

30

35

It is an object of the present invention to provide for enhanced replication-defective recombinant adenoviral vaccine vector backbones. These recombinant adenoviral backbones may accept one or more transgenes, which may be passaged through cell culture for growth, amplification and harvest.

It is a further object to provide for enhanced replication-defective recombinant adenoviral vaccine vectors which encode various transgenes.

It is also an object of the present invention to provide for a harvested recombinant, replication-deficient adenovirus which shows enhanced growth and amplification rates while in combination with increased virus stability after continuous passage in cell culture. Such a recombinant adenovirus is particularly suited for use in gene therapy and nucleotide-based vaccine vectors which, favorably, lends itself to large scale propagation.

To this end, it is an object of the present invention to provide for (1) enhanced replication-defective recombinant adenoviral vaccine vectors as described herein which encode various forms of HIV-1 Gag, HIV-1 Pol, and/or HIV-1 Nef, including immunologically relevant modifications of HIV-1 Gag, HIV-1 Pol and HIV-1 Nef, and (2) harvested, purified recombinant replication-deficient adenovirus generated by passage of the adenoviral vectors of (1) through one or multiple passages through cell culture, including but not limited to passage through 293 cells or PER.C6® cells.

It is also an object of the present invention to provide for recombinant adenovirus harvested by one or multiple passages through cell culture. As relating to recombinant adenoviral vaccine vector, this recombinant virus is harvested and formulated for subsequent host administration.

It is also an object of the present invention to provide for replication-defective adenoviral vectors wherein at least one gene is inserted in the form of a gene expression cassette comprising (a) a nucleic acid encoding a protein or biologically active and/or immunologically relevant portion thereof; (b) a heterologous promoter operatively linked to the nucleic acid of part a); and, (c) a transcription terminator.

5

10

15

20

25

30

35

It is also an object of the present invention to provide for a host cell comprising said adenoviral vectors and/or said shuttle plasmid vectors; vaccine compositions comprising said vectors; and methods of producing the vectors comprising (a) introducing the adenoviral vector into a host cell which expresses adenoviral E1 protein, and (b) harvesting the resultant adenoviral vectors. It is a further object of the present invention to provide for methods of generating a cellular immune response against a protein in an individual comprising administering to the individual an adenovirus vaccine vector comprising a) a replication defective adenoviral vector, at least partially deleted in E1, comprising a wildtype adenovirus cis-acting packaging region from about base pair 1 to between from about base pair 342 (more preferably, 400) to about 450 (preferably, 1-450) and, preferably, 3511-3523 of a wildtype adenovirus sequence, and, b) a gene expression cassette comprising:(i) a nucleic acid encoding a protein or biologically active and/or immunologically relevant portion thereof; (ii) a heterologous promoter operatively linked to the nucleic acid of part a); and (iii) a transcription terminator and/or a polyadenylation site.

It is also an object of the present invention to provide various alternatives for vaccine administration regimes, namely administration of one or more adenoviral and/or DNA plasmid vaccines described herein to provide effective immunoprophylaxis for uninfected individuals or a therapeutic treatment for HIV infected patients. Such processes include but are not limited to multivalent HIV-1 vaccine compositions, various combined modality regimes as well as various prime/boost alternatives. These methods of administration, relating to vaccine composition and/or scheduled administration, will increase the probability of eliciting an even more potent cellular immune response when compared to inoculation with a single modality regimen.

As used throughout the specification and claims, the following definitions and abbreviations are used:

"HAART" refers to -- highly active antiretroviral therapy --.

"first generation" vectors are characterized as being replication-defective.

They typically have a deleted or inactivated E1 gene region, and preferably have a deleted or inactivated E3 gene region as well.

"AEX" refers to Anion Exchange chromatography.

"QPA" refers to Quick PCR-based Potency Assay.

"bps" refers to basepairs.

"s" or "str" denotes that the transgene is in the E1 parallel or "straight" orientation.

"PBMCs" refers to peripheral blood monocyte cells.

"FL" refers to full length.

15

20

25

30

35

"FLgag" refers to a full-length optimized gag gene, as shown in Figure 2.

"Ad5-Flgag" refers to an adenovirus serotype 5 replication deficient virus which carries an expression cassette which comprises a full length optimized gag gene under the control of a CMV promoter.

"Promoter" means a recognition site on a DNA strand to which an RNA polymerase binds. The promoter forms an initiation complex with RNA polymerase to initiate and drive transcriptional activity. The complex can be modified by activating sequences such as enhancers or inhibiting sequences such as silencers.

"Leader" means a DNA sequence at the 5' end of a structural gene which is transcribed along with the gene. This usually results a protein having an N-terminal peptide extension, often referred to as a pro-sequences.

"Intron" means a section of DNA occurring in the middle of a gene which does not code for an amino acid in the gene product. The precursor RNA of the intron is excised and is therefore not transcribed into mRNA not translated into protein.

"Immunologically relevant" or "biologically active" means (1) with regards to a viral protein, that the protein is capable, upon administration, of eliciting a measurable immune response within an individual sufficient to retard the propagation and/or spread of the virus and/or to reduce the viral load present within the individual; or (2) with regards to a nucleotide sequence, that the sequence is capable of encoding for a protein capable of the above.

"Cassette" refers to a nucleic acid sequence which is to be expressed, along with its transcription and translational control sequences. By changing the cassette, a vector can express a different sequence.

"bGHpA" refers to the bovine growth hormone transcription terminator/polyadenylation sequence.

"tPAgag" refers to a fusion between the leader sequence of the tissue plasminogen activator leader sequence and an optimized HIV gag gene, as exemplified in Figure 30A-B, whether in a DNA or adenovirus-based vaccine vector.

Where utilized, "IA" or "inact" refers to an <u>inactivated</u> version of a gene (e.g. IApol).

"MCS" is "multiple cloning site".

5

10

15

20

25

30

35

In general, adenoviral constructs, gene constructs are named by reference to the genes contained therein. For example:

"Ad5 HIV-1 gag", also referred to as the original HIV-1 gag adenoviral vector, is a vector containing a transgene cassette composed of a hCMV intron A promoter, the full length version of the human codon-optimized HIV-1 gag gene, and the bovine growth hormone polyadenylation signal. The transgene was inserted in the E1 antiparallel orientation in an E1 and E3 deleted adenovector.

"MRK Ad5 HIV-1 gag" also referred to as "MRKAd5gag" or "Ad5gag2" is an adenoviral vector taught herein which is deleted of E1, comprises basepairs 1-450 and 3511-3523, and has a human codon-optimized HIV-1 gene in an E1 parallel orientation under the control of a CMV promoter without intron A. The construct also comprises a bovine growth hormone polyadenylation signal.

"pV1JnsHIVgag", also referred to as "HIVFLgagPR9901", is a plasmid comprising the CMV immediate-early (IE) promoter and intron A, a full-length codon-optimized HIV gag gene, a bovine growth hormone-derived polyadenylation and transcriptional termination sequence, and a minimal pUC backbone.

"pV1JnsCMV(no intron)-FLgag-bGHpA" is a plasmid derived from pV1JnsHIVgag which is deleted of the intron A portion of CMV and which comprises the full length HIV gag gene. This plasmid is also referred to as "pV1JnsHIVgag-bGHpA", pV1Jns-hCMV-FL-gag-bGHpA" and "pV1JnsCMV(no intron) + FLgag + bGHpA".

"pV1JnsCMV(no intron)-FLgag-SPA" is a plasmid of the same composition as pV1JnsCMV(no intron)-FLgag-bGHpA except that the SPA termination sequence replaces that of bGHpA. This plasmid is also referred to as "pV1Jns-HIVgag-SPA" and pV1Jns-hCMV-FLgag-SPA".

"pdelE1sp1A" is a universal shuttle vector with no expression cassette (i.e., no promoter or polyA). The vector comprises wildtype adenovirus serotype 5 (Ad5) sequences from bp 1 to bp 341 and bp 3524 to bp 5798, and has a multiple cloning

site between the Ad5 sequences ending 341 bp and beginning 3524 bp. This plasmid is also referred to as the original Ad 5 shuttle vector.

"MRKpdelE1sp1A" or "MRKpdelE1(Pac/pIX/pack450)" or

5

10

15

20

25

30

35

"MRKpdelE1(Pac/pIX/pack450)Cla1" is a universal shuttle vector with no expression cassette (i.e. no promoter or polyA) comprising wildtype adenovirus serotype 5 (Ad5) sequences from bp1 to bp450 and bp 3511 to bp 5798. The vector has a multiple cloning site between the Ad5 sequence ending 450 bp and beginning 3511 bp. This shuttle vector may be used to insert the CMV promoter and the bGHpA fragments in both the straight ("str". or E1 parallel) orientation or in the opposite (opp. or E1 antiparallel) orientation)

"MRKpdelE1(Pac/pIX/pack450)+CMVmin+BGHpA(str.)" is still another shuttle vector which is the modified vector that contains the CMV promoter (no intronA) and the bGHpA fragments. The expression unit containing the hCMV promoter (no intron A) and the bovine growth hormone polyadenylation signal has been inserted into the shuttle vector such that insertion of the gene of choice at a unique *BgI*II site will ensure the direction of transcription of the transgene will be Ad5 E1 parallel when inserted into the MRKpAd5(E1/E3+)Cla1 pre-plasmid. This shuttle vector, as shown in Figures 22 and 23, was used to insert the respective IApol and G2A,LLAA nef genes directly into.

"MRKpdelE1-CMV(no intron)-FLgag-bGHpA" is a shuttle comprising Ad5 sequences from basepairs 1-450 and 3511-5798, with an expression cassette containing human CMV without intron A, the full-length human codon-optimized HIV gag gene and bovine growth hormone polyadenylation signal. This plasmid is also referred to as "MRKpdelE1 shuttle +hCMV-FL-gag-BGHpA"

"MRKpAdHVE3+CMV(no intron)-FLgag-bGHpA" is an adenoviral vector comprising all Ad5 sequences except those nucleotides encompassing the E1 region (from 451-3510), a human CMV promoter without intron A, a full-length human codon-optimized HIV gag gene, and a bovine growth hormone polyadenylation signal. This vector is also referred to as "MRKpAdHVE3 + hCMV-FL-gag-BGHpA", "MRKpAd5HIV-1gag", "MRKpAd5gag", "pMRKAd5gag" or "pAd5gag2".

"pV1Jns-HIV-pol inact(opt)" or "pV1Jns-HIV IA pol (opt) is the inactivated Pol gene (contained within SEQ ID NO:3) cloned into the BglII site of V1Jns (Figure 17A-C). As noted herein, various derivatives of HIV-1 pol may be cloned into a plasmid expression vector such as V1Jns or V1Jns-tPA, thus serving directly as DNA vaccine candidates or as a source for subcloning into an appropriate adenoviral vector.

"MRKpdel+hCMVmin+FL-pol+bGHpA(s)" is the "MRKpdelE1(Pac/pIX/pack450)+CMVmin+BGHpA(str.)" shuttle mentioned above which contains the IA pol gene is the proper orientation. This shuttle vector is used in a bacterial recombination with MRKpAd(E1-/E3+)Cla1.

"MRKpAd+hCMVmin+FL-pol+bGHpA(S)E3+", also referred to herein as "pMRKAd5pol", is the pre-adenovirus plasmid which comprises a CMV-pol inact(opt)-pGHpA construct. The construction of this pre-adenovirus plasmid is shown in Figure 22.

"pV1Jns/nef (G2A,LLAA)" or "V1Jns/opt nef (G2A,LLAA)" comprises codon optimized HTV-1 Nef wherein the open reading frame codes for modifications at the amino terminal myristylation site (Gly-2 to Ala-2) and substitution of the Leu-174-Leu-175 dileucine motif to Ala-174-Ala-175 (SEQ ID NO:13; which comprises an initiating methionine residue at nucleotides 12-14 and a "TAA" stop codon from nucleotides 660-662). This fragment is subcloned into the Bgl II site of V1Jns and/orV1Jns-tPA (Figures 16A-B). As noted above for HIV-1 pol, HIV-1 nef constructs may be cloned into a plasmid expression vector such as V1Jns or V1Jns-tPA, thus serving directly as DNA vaccine candidates or as a source for subcloning into an appropriate adenoviral vector.

"MRKpdelE1hCMVminFL-nefBGHpA(s)", also referred to herein as "pMRKAd5nef", is the pre-adenovirus plasmid which comprises a CMV-nef (G2A,LLAA) codon optimized sequence. The construction of this pre-adenovirus plasmid is shown in Figure 23.

### BRIEF DESCRIPTION OF THE FIGURES

5

10

15

20

25

30

35

Figure 1 shows the original HIV-1 gag adenovector (Ad5HIV-1gag). This vector is disclosed in PCT International Application No. PCT/US00/18332 (WO 01/02607) filed July 3, 2000, claiming priority to U.S. Provisional Application Serial No. 60/142,631, filed July 6, 1999 and U.S. Application Serial No. 60/148,981, filed August 13, 1999, all three applications which are hereby incorporated by reference.

Figure 2 shows the nucleic acid sequence (SEQ ID NO: 29) of the optimized human HIV-1 gag open reading frame.

Figure 3 shows diagrammatically the new transgene constructs in comparison with the original gag transgene.

Figure 4 shows the modifications made to the original adenovector backbone in the generation of the novel vectors of the instant invention.

Figure 5 shows the virus mixing experiments that were carried out to determine the effects of the addition made to the packaging signal region (Expt. #1) and the E3 gene on viral growth (Expt. #2). The bars denote the region of modifications made to the E1 deletion.

Figure 6 shows an autoradiograph of viral DNA analysis following the viral mixing experiments described in Examples 6 and 7.

5

10

15

20

25

30

35

Figures 7A, 7B and 7C are as follows: Figure 7A shows the hCMV-Flgag-bGHpA adenovectors constructed within the MRKpAdHVE3 and MRKpAdHVO adenovector backbones. Both E1 parallel and E1 antiparallel transgene orientation are represented. Figure 7B shows the hCMV-Flgag-SPA adenovectors constructed within the MRKpAdHVE3 and MRKpAdHVO adenovector backbones. Again, both E1 parallel and E1 antiparallel transgene orientation are represented. Figure 7C shows the mCMV-Flgag-bGHpA adenovectors constructed within the MRKpAdHVE3 and MRKpAdHVO adenovector backbones. Once again, both E1 parallel and E1 antiparallel transgene orientation are represented.

Figure 8A shows the experiment designed to test the effect of transgene orientation.

Figure 8B shows the experiments designed to test the effect of polyadenylation signal.

Figure 9 shows viral DNA from the four adenoviral vectors tested (Example 12) at P5, following *Bst*E11 digestion.

Figure 10 shows viral DNA analysis of passages 11 and 12 of MRKpAdHVE3, MRKAd5HIV-1gag, and MRKAd5HIV-1gagE3-.

Figure 11 shows viral DNA analysis (*Hind*III digestion) of passage 6 MRKpAdHVE3 and MRKAd5HIV-1gag used to initiate the viral competition study. The last two lanes are passage 11 analysis of duplicate passages of the competition study (each virus at MOI of 280 viral particles).

Figure 12 shows viral DNA analysis by *Hind* III digestion on high passage numbers for MRKAd5HIV-1gag in serum-containing media with collections made at specified times. The first lane shows the 1kb DNA size marker. The other lanes represent pre-plasmid control (digested with Pac1 and *Hind*III), MRKAd5HIV-1gag at P16, P19, and P21.

Figure 13 shows serum anti-p24 levels at 3 wks post i.m. immunization of balb/c mice (n=10) with varying doses of several Adgag constructs: (A) MRK Ad5 HIV-1 gag (through passage 5); (B) MRKAd5 hCMV-FLgag-bGHpA (E3-); (C) MRKAd5 hCMV-FLgag-SPA (E3+); (D) MRKAd5 mCMV-FLgag-bGHpA (E3+);

(E) research lot (293 cell-derived) of Ad5HIV-1 gag; and (F) clinical lot (Ad5gagFN0001) of Ad5HIV-1 gag. Reported are the geometric mean titers (GMT) for each cohort along with the standard error bars.

Figure 14 shows a restriction map of the pMRKAd5HIV-1gag vector.

5

10

15

20

25

30

35

Figures 15A-X illustrates the nucleotide sequence of the pMRKAd5HIV-1gag vector (SEQ ID NO:27.[coding] and SEQ ID NO:28 [non-coding]).

Figures 16A-B shows a schematic representation of DNA vaccine expression vectors V1Jns (A) and V1Jns-tPA (B), which are utilized for HIV-1 gag, pol and nef constructs in various DNA/viral vector combined modality regimens as disclosed herein.

Figures 17A-C shows the nucleotide (SEQ ID NO:3) and amino acid sequence (SEQ ID NO:4) of IA-Pol. Underlined codons and amino acids denote mutations, as listed in Table 1.

Figure 18 shows codon optimized nucleotide and amino acid sequences through the fusion junction of tPA-pol inact(opt) (contained within SEQ ID NOs: 7 and 8, respectively). The underlined portion represents the NH<sub>2</sub>-terminal region of IA-Pol.

Figures 19A-B show a nucleotide sequence comparison between wild type nef(jrfl) and codon optimized nef. The wild type nef gene from the jrfl isolate consists of 648 nucleotides capable of encoding a 216 amino acid polypeptide. WT, wild type sequence (SEQ ID NO:19); opt, codon-optimized sequence (contained within SEQ ID NO:1). The Nef amino acid sequence is shown in one-letter code (SEQ ID NO:2).

Figures 20A-C show nucleotide sequences at junctions between nef coding sequence and plasmid backbone of nef expression vectors V1Jns/nef (Figure 20A), V1Jns/nef(G2A,LLAA) (Figure 20B), V1Jns/tpanef (Figure 20C) and V1Jns/tpanef(LLAA) (Figure 20C, also). 5' and 3' flanking sequences of codon optimized nef or codon optimized nef mutant genes are indicated by bold/italic letters; nef and nef mutant coding sequences are indicated by plain letters. Also indicated (as underlined) are the restriction endonuclease sites involved in construction of respective nef expression vectors. V1Jns/tpanef and V1Jns/tpanef(LLAA) have identical sequences at the junctions.

Figure 21 shows a schematic presentation of nef and nef derivatives. Amino acid residues involved in Nef derivatives are presented. Glycine 2 and Leucine 174 and 175 are the sites involved in myristylation and dileucine motif, respectively. For both versions of the tpanef fusion genes, the putative leader peptide cleavage sites are

indicated with "\*", and a exogenous serine residue introduced during the construction of the mutants is underlined.

Figure 22 shows diagrammatically the construction of the pre-adenovirus plasmid construct, MRKAd5Pol.

Figure 23 shows diagrammatically the construction of the pre-adenovirus plasmid construct, MRKAd5Nef.

5

10

15

20

25

30

35

Figure 24 shows a comparison of clade B vs. clade C anti-gag T cell responses in clade B HIV-infected subjects.

Figure 25 shows a comparison of clade B vs. clade C anti-nef T cell responses in clade B HIV-infected subjects.

Figures 26A-AO illustrates the nucleotide sequence of the pMRKAd5HIV-1pol adenoviral vector (SEQ ID NO:32 [coding] and SEQ ID NO:33 [non-coding]), comprising the coding region of the inactivated pol gene (SEQ ID NO3).

Figures 27A-AM illustrates the nucleotide sequence of the pMRKAd5HIV-1 nef adenoviral vector (SEQ ID NO:34 [coding] and SEQ ID NO:35 [non-coding]), comprising the coding region of the inactivated pol gene (SEQ ID NO13).

Figure 28 shows the stability of MRKAd5 vectors comprising various promoter fragments (hCMV or mCMV) and terminations signals (bGH or SPA) in E3(+) or E3(-) backbones.

Figures 29A and B shows the anion-exchange HPLC viral particle concentrations of the freeze-thaw recovered cell associated virus at the 24, 36, 48, and 60 hpi time points (Figure 29A) and the timcourse QPA supernatant titers (Figure 29B) for MRKAd5gag, MRKAd5pol and MRKAd5nef.

Figure 30 shows the nucleotide sequence (SEQ ID NO:36) and amino acid sequence (SEQ ID NO:37) comprising the open reading frame of a representative tPA-gag fusion for use in the DNA and/or adenoviral vaccine disclosed herein.

Figure 31 shows the intracellular γIFN staining of PBMCs collected at week 10 (post DNA prime) and week 30 (post Ad boost). The cells were stimulated overnight in the presence or absence of the gag peptide pool. They were subsequently stained using fluorescence-tagged anti-CD3, anti-CD8, anti-CD4, and anti-γIFN monoclonal antibodies. Each plot shows all CD3+ T cells which were segregated in terms of positive staining for surface CD8 and γIFN production. The numbers in the upper right and lower right quadrants of each plot are the percentages of CD3<sup>+</sup> cells that were CD8<sup>+</sup>γIFN<sup>+</sup> and CD4<sup>+</sup>γIFN<sup>+</sup>, respectively.

Figure 32 shows a comparison of single-modality adenovirus immunization with DNA + adjuvant prime/adenovirus boost immunization.

Figures 33A-B show the nucleotide sequence (SEQ ID NO: 38) of the open reading frame for the gag-IApol fusion of Example 29.

Figures 34A-B show the protein sequence (SEQ ID NO:39) of the gag-IApol fustion frame.

5

10

15

20

25

30

35

### DETAILED DESCRIPTION OF THE INVENTION

A novel replication-defective, or "first generation," adenoviral vector suitable for use in gene therapy or nucleotide-based vaccine vectors is described. This vector is at least partially deleted in E1 and comprises a wildtype adenovirus cis-acting packaging region from about base pair 1 to between about base pair 342 (more preferably, 400) to about 458 (preferably, 1-450) and, preferably, 3511-3523 of a wild-type adenovirus sequence. It has been found that a vector of this description possesses enhanced growth characteristics, with approximately 5-10 fold greater amplification rates, and is more potent allowing lower doses of virus to be used to generate equivalent immunity. The vector, furthermore, generates a harvested recombinant adenovirus which shows greater cellular-mediated immune responses than replication-deficient vectors not comprising this region (basepairs 342-450). Adenoviral constructs derived from these vectors are, further, very stable genetically, particularly those comprising a transgene under the control of a hCMV promoter devoid of intron A. Viruses in accordance with this description were passaged continually and analyzed; see Example 12. Each virus analyzed maintained it correct genetic structure. Analysis was also carried out under propagation conditions similar to that performed in large scale production. Again, the vectors were found to possess enhanced genetic stability; see Figure 12. Following 21 passages, the viral DNA showed no evidence of rearrangement, and was highly reproducible from one production lot to the next. The outcome of all relevant tests indicate that the adenoviral vector is extremely well suited for large-scale production of recombinant, replication-deficient adenovirus, as shown herein with the data associated with Figure 28.

A preferred adenoviral vector in accordance with this description is a vector comprising basepairs 1-450, which is deleted in E3. This vector can accommodate up to approximately 7,500 base pairs of foreign DNA inserts (or exogenous genetic material). Another preferred vector is one retaining E3 which comprises basepairs 1-450. A preferred vector of this description is an E3+ vector comprising basepairs 1-450 and 3511-3523. This vector, when deleted of the region spanning basepairs 451-3510, can accommodate up to approximately, 4,850 base pairs of foreign DNA inserts

(or exogenous genetic material). The cloning capacities of the above vectors have been determined using 105% of the wildtype Ad5 sequence as the upper genome size limit.

Wildtype adenovirus serotype 5 is used as the basis for the specific basepair numbers provided throughout the specification. The wildtype adenovirus serotype 5 sequence is known and described in the art; see, Chroboczek et al., 1992 J. Virology 186:280, which is hereby incorporated by reference. Accordingly, a particular embodiment of the instant invention is a vector based on the adenovirus serotype 5 sequence. One of skill in the art can readily identify the above regions in other adenovirus serotypes (e.g., serotypes 2, 4, 6, 12, 16, 17, 24, 31, 33, and 42), regions defined by basepairs corresponding to the above basepair positions given for adenovirus serotype 5. Accordingly, the instant invention encompasses all adenoviral vectors partially deleted in E1 comprising basepairs corresponding to 1-450 (particularly, 342-450) and, preferably, 3511-3523 of a wild-type adenovirus serotype 5 (Ad5) nucleic acid sequence. Particularly preferred embodiments of the instant invention are those derived from adenoviruses like Ad5 which are classified in subgroup C (e.g., Ad2).

10

15

20

25

30

35

Vectors in accordance with the instant invention are at least partially deleted in E1. Preferably the E1 region is completely deleted or inactivated. Most preferably, the region deleted of E1 is within basepairs 451-3510. It is to be noted that the extended 5' and 3' regions of the disclosed vectors are believed to effectively reduce the size of the E1 deletion of previous constructs without overlapping any part of the E1A/E1B gene present in the cell line used, i.e., the PER.C6® cell line transfected with base pairs 459-3510. Overlap of adenoviral sequences is avoided because of the possibility of recombination. One of ordinary skill in the art can certainly appreciate that the instant invention can, therefore, be modified if a different cell line transfected with a different segment of adenovirus DNA is utilized. For purposes of exemplification, a 5' region of base pairs 1 to up to 449 is more appropriate if a cell line is transfected with adenoviral sequence from base pairs 450-3510. This holds true as well in the consideration of segments 3' to the E1 deletion.

Preferred embodiments of the instant invention possess an intact E3 region (i.e., an E3 gene capable of encoding a functional E3). Alternate embodiments have a partially deleted E3, an inactivated E3 region, or a sequence completely deleted of E3. Applicants have found, in accordance with the instant invention, that virus comprising the E3 gene were able to amplify more rapidly compared with virus not comprising an E3 gene; see Figure 6 wherein a diagnostic CsCl band corresponding to the E3+ virus

tested (5,665 bp) was present in greater amount compared with the diagnostic band of 3,010 bp corresponding to the E3- virus. These results were obtained following a virus competition study involving mixing equal MOI ratio (1:1) of adenovectors both comprising the E3 gene and not comprising the E3 gene. This increased amplification capacity of the E3+ adenovectors was subsequently confirmed with growth studies; see Table 4A, wherein the E3+ virus exhibit amplification ratios of 470, 420 and 320 as compared with the 115 and 40-50 of the E3- constructs.

5

10

15

20

25

30

35

As stated above, vectors in accordance with the instant invention can accommodate up to approximately 4,850 base pairs of exogenous genetic material for an E3+ vector and approximately 7,500 base pairs for an E3- vector. Preferably, the insert brings the adenoviral vector as close as possible to a wild-type genomic size (e.g., for Ad5, 35,935 basepairs). It is well known that adenovirus amplifies best when they are close to their wild-type genomic size.

The genetic material can be inserted in an E1-parallel or an E1 anti-parallel orientation, as such is illustrated in Figure 7A, 7B, 7C and Figure 8A. Particularly preferred embodiments of the instant invention, have the insert in an E1-parallel orientation. Applicants have found, via competition experiments with plasmids containing transgenes in differing orientation (Figure 8A), that vector constructs with the foreign DNA insert in an E1-parallel orientation amplify better and actually outcompete E1-antiparallel-oriented transgenes. Viral DNA analysis of the mixtures at passage 3 and certainly at passage 6, showed a greater ratio of the virus carrying the transgene in the E1 parallel orientation as compared with the E1 anti-parallel version. By passage 10, the only viral species observed was the adenovector with the transgene in the E1 parallel orientation for both transgenes tested.

Adenoviral vectors in accordance with the instant invention are particularly well suited to effectuate expression of desired proteins, one example of which is an HIV protein, particularly an HIV full length gag protein. Exogenous genetic material encoding a protein of interest can exist in the form of an expression cassette. A gene expression cassette preferably comprises (a) a nucleic acid encoding a protein of interest; (b) a heterologous promoter operatively linked to the nucleic acid encoding the protein; and (c) a transcription terminator.

The transcriptional promoter is preferably recognized by an eukaryotic RNA polymerase. In a preferred embodiment, the promoter is a "strong" or "efficient" promoter. An example of a strong promoter is the immediate early human cytomegalovirus promoter (Chapman et al, 1991 *Nucl. Acids Res*19:3979-3986, which is incorporated by reference), preferably without intronic sequences. Most preferred

for use within the instant adenoviral vector is a human CMV promoter without intronic sequences, like intron A. Applicants have found that intron A, a portion of the human cytomegalovirus promoter (hCMV), constitutes a region of instability for adenoviral vectors. CMV without intron A has been found to effectuate (Examples 1-3) comparable expression capabilities in vitro when driving HIV gag expression and, furthermore, behaved equivalently to intron A-containing constructs in Balb/c mice in vivo with respect to their antibody and T-cell responses at both dosages of plasmid DNA tested (20 µg and 200 µg). Those skilled in the art will appreciate that any of a number of other known promoters, such as the strong immunoglobulin, or other eukaryotic gene promoters may also be used, including the EF1 alpha promoter, the murine CMV promoter, Rous sarcoma virus (RSV) promoter, SV40 early/late promoters and the beta-actin promoter.

5

10

15

20

25

30

35

In preferred embodiments, the promoter may also comprise a regulatable sequence such as the Tet operator sequence. This would be extremely useful, for example, in cases where the gene products are effecting a result other than that desired and repression is sought.

The combination of the CMV promoter (devoid of the intron A region) with the BGH terminator is particularly preferred although other promoter/terminator combinations in the context of FG adenovirus may also be used.

Other embodiments incorporate a leader or signal peptide into the transgene. A preferred leader is that from the tissue-specific plasminogen activator protein, tPA. Examples include but are not limited to the various tPA-gag, tPA-pol and tPA-nef adenovirus-based vaccines disclosed throughout this specification.

In view of the improved adenovirus vectors described herein, an essential portion of the present invention are adenoviral-based HIV vaccines comprising said adenovirus backbones which may be administered to a mammalian host, preferably a human host, in either a prophylactic or therapeutic setting. The HIV vaccines of the present invention, whether administered alone or in combination regimens with other viral- or non-viral-based DNA vaccines, should elicit potent and broad cellular immune responses against HIV that will either lessen the likelihood of persistent virus infection and/or lead to the establishment of a clinically significant lowered virus load

subject to HIV infection or in combination with HAART therapy, mitigate the effects of previously established HIV infection (antiviral immunotherapy(ARI)). While any HIV antigen (e.g., gag, pol, nef, gp160, gp41, gp120, tat, rev, etc.) may be utilized in the herein described recombinant adenoviral vectors, preferred embodiments include the codon optimized p55 gag antigen (herein exemplified as MRKAd5gag), pol and nef. Sequences based on different Clades of HIV-1 are suitable for use in the instant invention, most preferred of which are Clade B and Clade C. Particularly preferred embodiments are those sequences (especially, codon-optimized sequences) based on concensus Clade B sequences. Preferred versions of the MRKAd5pol and MRKAd5nef series of adenoviral vaccines will encode modified versions of pol or nef, as discussed herein. Preferred embodiments of the MRKAd5HIV-1 vectors carrying HIV envelope genes and modifications thereof comprise the HIV codon-optimized env sequences of PCT International Applications PCT/US97/02294 and PCT/US97/10517, published August 28, 1997 (WO 97/31115) and December 24, 1997, respectively; both documents of which are hereby incorporated by reference.

5

10

15

20

25

30

35

A most preferred aspect of the instant invention is the disclosed use of the adenoviral vector described above to effectuate expression of HIV gag. Sequences for many genes of many HIV strains are publicly available in GENBANK and primary, field isolates of HIV are available from the National Institute of Allergy and Infectious Diseases (NIAID) which has contracted with Quality Biological (Gaithersburg, MD) to make these strains available. Strains are also available from the World Health Organization (WHO), Geneva Switzerland. It is preferred that the gag gene be from an HIV-1 strain (CAM-1; Myers et al, eds. "Human Retroviruses and AIDS: 1995, IIA3-IIA19, which is hereby incorporated by reference). This gene closely resembles the consensus amino acid sequence for the clade B (North American/European) sequence. Therefore, it is within the purview of the skilled artisan to choose an appropriate nucleotide sequence which encodes a specific HIV gag antigen, or immunologically relevant portion thereof. As shown in Example 25, a clade B or clade C based p55 gag antigen will potentially be useful on a global scale. As noted herein, the transgene of choice for insertion in to a DNA or MRKAd-based adenoviral vector of the present invention is a codon optimized version of p55 gag. Such a MRKAd5gag adenoviral vector is documented in Example 11 and is at least referred to herein as MRKAd5HIV-1gag. Of course, additional versions are contemplated, including but not limited to modifications such as promoter (e.g., mCMV for hCMV) and/or pA-terminations signal (SPA for bGH) switching, as well as generating MRK Ad5 backbones with or without deletion of the Ad5 E3 gene.

5

10

15

20

25

30

35

The present invention also relates a series of MRKAd5pol-based adenoviral vaccines which are shown herein to generate cellular immune responses subsequent to administration in mice and non-human primate studies. Several of the MRKAd5pol series are exemplified herein. One such adenoviral vector is referred to as MRKAd5hCMV-inact opt pol(E3+), which comprises the MRKAd5 backbone, the hCMV promoter (no intron A), an inactivated pol transgene, and contains the Ad5 E3 gene in the adenoviral backbone. A second exemplified pre-adenovirus plasmid and concomitant virus is referred to as MRKAd5hCMV-inact opt pol(E3-), which is identical to the former adenoviral vector except that the E3 is deleted. Both constructions contain a codon optimized, inactivated version of HIV-1 Pol, wherein at least the entire coding region is disclosed herein as SEQ ID NO:3 and the expressed protein is shown as SEQ ID NO:4 (see also Figure 17A-C and Table 1, which show targeted deletion for inactivated pol. This and other preferred codon optimized versions of HIV Pol as disclosed herein are essentially as described in U.S. Application Serial No. 09/745,221, filed December 21, 2000 and PCT International Application PCT/US00/34724, also filed December 21, 2000, both documents which are hereby incorporated by reference. As disclosed in the above-mentioned documents, the open reading frame for these codon-optimized HIV-1 Pol-based DNA vaccines are represented by codon optimized DNA molecules encoding codon optimized HIV-1 Pol (e.g. SEQ ID NO:2), codon optimized HIV-1 Pol fused to an amino terminal localized leader sequence (e.g. SEQ ID NO:6), and especially preferable, and exemplified by the MRKAd5-Pol construct in e.g., Example 19. biologically inactivated pol ("inact opt Pol"; e.g., SEQ ID NO:4) which is devoid of significant PR, RT, RNase or IN activity associated with wild type Pol. In addition, a construct related to SEQ ID NO:4 is contemplated which contains a leader peptide at the amino terminal region of the IA Pol protein. A specific construct is ligated within an appropriate DNA plasmid vector containing regulatory regions operatively linked to the respective HIV-1 Pol coding region, with or without a nucleotide sequence encoding a functional leader peptide. To this end, various HIV-1 Pol constructs disclosed herein relate to open reading frames for cloning to the enhanced first generation Ad vectors of the present invention (such a series of MRKAd5pol adenoviral vaccine vectors), including but not limited to wild type Pol (comprising the DNA molecule encoding WT opt Pol, as set forth in SEQ ID NO:2), tPA-opt WTPol, (comprising the DNA molecule encoding tPA Pol, as set forth in SEQ ID NO:6), inact opt Pol (comprising the DNA molecule encoding IA Pol, as set forth in SEQ ID NO:4), and tPA-inact opt Pol, (comprising the DNA molecule encoding tPA-inact opt

Pol, as set forth in SEQ ID NO:8). The pol-based versions of enhanced first generation adenovirus vaccines elicit CTL and Th cellular immune responses upon administration to the host, including primates and especially humans. As noted in the above, an effect of the cellular immune-directed vaccines of the present invention should be a lower transmission rate to previously uninfected individuals and/or reduction in the levels of the viral loads within an infected individual, so as to prolong the asymptomatic phase of HIV-1 infection.

5

The present invention further relates to a series of MRKAd5nef-based adenoviral vaccines which, similar to HIV gag and pol antigens, generate cellular immune responses subsequent to administration in mice and non-human primate 10 studies. The MRKAd5nef series are exemplified herein by utilizing the improved MRK adenoviral backbone in combination with modified versions of HIV nef. These exemplified MRKAd5nef vectors are as follows: (1) MRKAd5hCMVnef(G2A,LLAA) (E3+), which comprises the improved MRKAd5 backbone, a human CMV promoter an intact Ad5 E3 gene and a modified nef gene: (2) MRKAd5mCMV-15 nef(G2A,LLAA) (E3+), which is the same as (1) above but substituting a murine CMV promoter for a human CMV promoter; and (3) MRKAd5mCMV-tpanef(LLAA) (E3+), which is the same as (2) except that the nef transgene is tpanef(LLAA). Codon optimized versions of HIV-1 Nef and HIV-1 Nef modifications are essentially as described in U.S. Application Serial No. 09/738,782, filed December 15, 2000 and 20 PCT International Application PCT/US00/34162, also filed December 15, 2000, both documents which are hereby incorporated by reference. Particular embodiments of codon optimized Nef and Nef modifications relate to a DNA molecule encoding HIV-1 Nef from the HIV-1 ifrl isolate wherein the codons are optimized for expression in a mammalian system such as a human. The DNA molecule which encodes this protein 25 is disclosed herein as SEQ ID NO:9, while the expressed open reading frame is disclosed herein as SEQ ID NO:10. Another embodiment of Nef-based coding regions for use in the adenoviral vectors of the present invention comprise a codon optimized DNA molecule encoding a protein containing the human plasminogen activator (tpa) leader peptide fused with the NH2-terminus of the HIV-1 Nef 30 polypeptide. The DNA molecule which encodes this protein is disclosed herein as SEQ ID NO:11, while the expressed open reading frame is disclosed herein as SEQ ID NO:12. Another modified Nef optimized coding region relates to a DNA molecule encoding optimized HIV-1 Nef wherein the open reading frame codes for modifications at the amino terminal myristylation site (Gly-2 to Ala-2) and 35 substitution of the Leu-174-Leu-175 dileucine motif to Ala-174-Ala-175, herein

described as opt nef (G2A, LLAA). The DNA molecule which encodes this protein is disclosed herein as SEQ ID NO:13, while the expressed open reading frame is disclosed herein as SEQ ID NO:14. MRKAd5nef vectors (1) MRKAd5hCMV-nef(G2A,LLAA) (E3+) and (2) MRKAd5mCMV-nef(G2A,LLAA) (E3+) contain this transgene. An additional embodiment relates to a DNA molecule encoding optimized HIV-1 Nef wherein the amino terminal myristylation site and dileucine motif have been deleted, as well as comprising a tPA leader peptide. This DNA molecule, opt tpanef (LLAA), comprises an open reading frame which encodes a Nef protein containing a tPA leader sequence fused to amino acid residue 6-216 of HIV-1 Nef (jfrl), wherein Leu-174 and Leu-175 are substituted with Ala-174 and Ala-175, herein referred to as opt tpanef (LLAA) is disclosed herein as SEQ ID NO:15, while the expressed open reading frame is disclosed herein as SEQ ID NO:16. The MRKAd5nef vector "MRKAd5mCMV-tpanef(LLAA) (E3+)" contains this transgene.

10

15

20

25

30

35

Along with the improved MRKAd5gag adenovirus vaccine vector described herein, generation of a MRKAd5pol and MRKAd5nef adenovirus vector provide for enhanced HIV vaccine capabilities. Namely, the generation of this trio of adenoviral vaccine vectors, all shown to generate effective cellular immune responses subsequent to host administration, provide for the ability to administer these vaccine candidates not only alone, but preferably as part of a divalent (i.e., gag and nef, gag and pol, or pol and nef components) or a trivalent vaccine (i.e., gag, pol and nef components). Therefore, a preferred aspect of the present invention are vaccine formulations and associated methods of administration and concomitant generation of host cellular immune responses associated with formulating three separate series of MRKAd5based adenoviral vector vaccines. Of course, this MRKAd5 vaccine series based on distinct HIV antigens promotes expanded opportunities for formulation of a divalent or trivalent vaccine, or possibly administration of separate formulations of one or more monovalent or divalent formulations within a reasonable window of time. It is also within the scope of the present invention to embark on combined modality regimes which include multiple but distinct components from a specific antigen. An example, but certainly not a limitation, would be separate MRKAd5pol vectors, with one vaccine vector expressing wild type Pol (SEQ ID NO:2) and another MRKAd5pol vector expressing inactivated Pol (SEQ ID NO:6). Another example might be separate MRKAd5nef vectors, with one vaccine vector expressing the tPA/LLAA version of Nef (SEQ ID NO:16) and another MRKAd5nef vector expressing the G2A,LLAA modified version of Nef (SEQ ID NO:14). Therefore, the MRKAd5 adenoviral vectors of the present invention may be used in combination

with multiple, distinct HIV antigen classes. Each HIV antigen class is subject to sequence manipulation, thus providing for a multitude of potential vaccine combinations; and such combinations are within the scope of the present invention. The utilization of such combined modalities vaccine formulation and administration increase the probability of eliciting an even more potent cellular immune response when compared to inoculation with a single modality regimen.

5

10

15

20

25

30

35

The present invention also relates to application of a mono-, dual-, or trimodality administration regime of the MRKAd5gag, pol and nef adenoviral vaccine series in a prime/boost vaccination schedule. This prime/boost schedule may include any reasonable combination of the MRKAd5gag, pol and nef adenoviral vaccine series disclosed herein. In addition, a prime/boost regime may also involve other viral and/or non-viral DNA vaccines. A preferable addition to an adenoviral vaccine vector regime includes but is not limited to plasmid DNA vaccines, especially DNA plasmid vaccines that contain at least one of the codon optimized gag, pol and nef constructions, as disclosed herein.

Therefore, one aspect of this invention is the administration of the adenoviral vector containing the optimized gag gene in a prime/boost regiment in conjunction with a plasmid DNA encoding gag. To distinguish this plasmid from the adenoviralcontaining shuttle plasmids used in the construction of an adenovirus vector, this plasmid will be referred to as a "vaccine plasmid" or "DNA plasmid vaccine". Preferred vaccine plasmids for use in this administration protocol are disclosed in pending U.S. patent application 09/017,981, filed February 3, 1998 and WO98/34640, published August 13, 1998, both of which are hereby incorporated by reference. Briefly, the preferred vaccine plasmid is designated V1Jns-FLgag, which expresses the same codon-optimized gag gene as the adenoviral vectors of this invention (see Figure 2 for the nucleotide sequence of the exemplified optimized codon version of full length p55 gag). The vaccine plasmid backbone, designated V1Jns contains the CMV immediate-early (IE) promoter and intron A, a bovine growth hormone-derived polyadenylation and transcription termination sequence as the gene expression regulatory elements, and a minimal pUC backbone; see Montgomery et al., 1993, DNA Cell Biol. 12:777-783. The pUC sequence permits high levels of plasmid production in E. coli and has a neomycin resistance gene in place of an ampicillin resistance gene to provide selected growth in the presence of kanamycin. Alternatively, a vaccine plasmid which has the CMV promoter deleted of intron A can be used. Those of skill in the art will recognize that alternative vaccine plasmid

vectors may be easily substituted for these specific constructs, and this invention specifically envisions use of such alternative plasmid DNA vaccine vectors.

5

10

15

20

25

30

35

Another aspect of the present invention is a prime/boost regimen which includes a vaccine plasmid which encodes an HIV pol antigen, preferably a codon optimized form of pol and also preferably a vaccine plasmid which comprises a nucleotide sequence which encodes a Pol antigen selected from the group of Pol antigens as shown in SEQ ID NOs: 2, 4, 6 and 8. The variety of potential DNA plasmid vaccines which encode various biologically active forms of HIV-1 Pol. wherein administration, intracellular delivery and expression of the HIV-1 Pol gene of interest elicits a host CTL and Th response. The preferred synthetic DNA molecules of the present invention encode codon optimized wild type Pol (without Pro activity) and various codon optimized inactivated HIV-1 Pol proteins. The HIV-1 pol open reading disclosed herein are especially preferred for pharmaceutical uses, especially for human administration as delivered via a recombinant adenoviral vaccine, especially an enhanced first generation recombinant adenoviral vaccine as described herein. Several embodiments of this portion of the invention are provided in detail below, namely DNA molecules which comprise a HIV-1 pol open reading frame, whether encoding full length pol or a modification or fusion as described herein, wherein the codon usage has been optimized for expression in a mammal, especially a human. Again, these DNA sequences are positioned appropriately within a recombinant adenoviral vector, such as the exemplified recombinant adenoviral vector described herein, so as to promote expression of the respective HIV-1 Pol gene of interest, and subsequent to administration, elicit a host CTL and Th response. Again, these preferred, but in no way limiting, pol genes are as disclosed herein and essentially as described in U.S. Application Serial No. 09/745,221, filed December 21, 2000 and PCT International Application PCT/US00/34724, also filed December 21, 2000, both documents which are hereby incorporated by reference.

A third series of vaccine plasmids which are useful in a combined modality and/or prime/boost regimen are vaccine plasmids which encode an HIV nef antigen or biologically and/or immunologically relevant modification thereof. As noted elsewhere, preferred vaccine plasmids contain a codon optimized form of nef and also preferably comprise a nucleotide sequence which encodes a Nef antigen selected from the group of Nef antigens as shown in SEQ ID NOs: 10, 12, 14 and 16. These preferred nef coding regions are disclosed herein, as well as being described in U.S. Application Serial No. 09/738,782, filed December 15, 2000 and PCT International

Application PCT/US00/34162, also filed December 15, 2000, both documents which are hereby incorporated by reference.

5

10

15

20

25

30

35

Therefore, the adenoviral vaccines and plasmid DNA vaccines of this invention may be administered alone, or may be part of a prime and boost administration regimen. A mixed modality priming and booster inoculation scheme will result in an enhanced immune response, particularly is pre-existing anti-vector immune responses are present. This one aspect of this invention is a method of priming a subject with the plasmid vaccine by administering the plasmid vaccine at least one time, allowing a predetermined length of time to pass, and then boosting by administering the adenoviral vaccine. Multiple primings typically, 1-4, are usually employed, although more may be used. The length of time between priming and boost may typically vary from about four months to a year, but other time frames may be used. In experiments with rhesus monkeys, the animals were primed four times with plasmid vaccines, then were boosted 4 months later with the adenoviral vaccine. Their cellular immune response was notably higher than that of animals which had only received adenoviral vaccine. The use of a priming regimen may be particularly preferred in situations where a person has a pre-existing anti-adenovirus immune response.

Furthermore and in the alternative, multiple HIV-1 viral antigens, such as the MRKAd5 adenoviral vaccines disclosed herein, may be ligated into a proper shuttle plasmid for generation of a pre-adenoviral plasmid comprising multiple open reading frames. For example a trivalent vector may comprise a gag-pol-nef fusion, in either a E3(-) or E3(+) background, preferably a E3 deleted backbone, or possible a "2+1" divalent vaccine, such as a gag-pol fusion (i.e., codon optimized p55 gag and inactivated optimized pol; Example 29 and Table 25) within the same MRKAd5 backbone, with each open reading frame being operatively linked to a distinct promoter and transcription termination sequence. Alternatively, the two open reading frames may be operatively linked to a single promoter, with the open reading frames operatively linked by an internal ribosome entry sequence (IRES), as disclosed in International Publication No. WO 95/24485, which is hereby incorporated by reference. Figure 9 shows that the use of multiple promoters and termination sequences provide for similar growth properties, while Figure 28 shows that these MRKAd5gag-based vectors are also stable at least through passage 21. In the absence of the use of IRES-based technology, it is preferred that a distinct promoter be used to support each respective open reading frame, so as to best preserve vector stability. As examples, and certainly not as limitations, potential multiple transgene vaccines may

include a three transgene vector such as hCMV-gagpol-bGHpA + mCMV-nef-SPA in an E3 deleted backbone or hCMV-gagpol-bGHpA + mCMV-nef-SPA(E3+). Potential "2+1" divalent vaccines of the present invention might be a hCMV-gagbGHpA + mCMV-nef-SPA in an E3+ backbone (vector #1) in combination with hCMV-pol-bGHpA in an E3+ backbone (vector #2), with all transgenes in the E1 parallel orientation. Fusion constructs other than the gag-pol fusion described above are also suitable for use in various divalent vaccine strategies and can be composed of any two HIV antigens fused to one another (e.g., nef-pol and gag-nef). These adenoviral compositions are, as above, preferably delivered along with an adenoviral composition comprising an additional HIV antigen in order to diversify the immune response generated upon administration. Therefore, a multivalent vaccine delivered in a single, or possible second, adenoviral vector is certainly contemplated as part of the present invention. Again, this mode of administration is another example of whereby an efficaceous adenovirus-based HIV-1 vaccine may be administered via a combined modality regime. It is important to note, however, that in terms of deciding on an insert for the disclosed adenoviral vectors, due consideration must be dedicated to the effective packaging limitations of the adenovirus vehicle. Adenovirus has been shown to exhibit an upper cloning capacity limit of approximately 105% of the wildtype Ad5 sequence.

5

10

15

20

25

30

35

Regardless of the gene chosen for expression, it is preferred that the sequence be "optimized" for expression in a human cellular environment. A "triplet" codon of four possible nucleotide bases can exist in 64 variant forms. That these forms provide the message for only 20 different amino acids (as well as transcription initiation and termination) means that some amino acids can be coded for by more than one codon. Indeed, some amino acids have as many as six "redundant", alternative codons while some others have a single, required codon. For reasons not completely understood. alternative codons are not at all uniformly present in the endogenous DNA of differing types of cells and there appears to exist variable natural hierarchy or "preference" for certain codons in certain types of cells. As one example, the amino acid leucine is specified by any of six DNA codons including CTA, CTC, CTG, CTT, TTA, and TTG (which correspond, respectively, to the mRNA codons, CUA, CUC, CUG, CUU, UUA and UUG). Exhaustive analysis of genome codon frequencies for microorganisms has revealed endogenous DNA of E. coli most commonly contains the CTG leucine-specifying codon, while the DNA of yeasts and slime molds most commonly includes a TTA leucine-specifying codon. In view of this hierarchy, it is generally held that the likelihood of obtaining high levels of expression of a leucine-

rich polypeptide by an *E. coli* host will depend to some extent on the frequency of codon use. For example, a gene rich in TTA codons will in all probability be poorly expressed in *E. coli*, whereas a CTG rich gene will probably highly express the polypeptide. Similarly, when yeast cells are the projected transformation host cells for expression of a leucine-rich polypeptide, a preferred codon for use in an inserted DNA would be TTA.

5

10

15

20

25

30

35

The implications of codon preference phenomena on recombinant DNA techniques are manifest, and the phenomenon may serve to explain many prior failures to achieve high expression levels of exogenous genes in successfully transformed host organisms—a less "preferred" codon may be repeatedly present in the inserted gene and the host cell machinery for expression may not operate as efficiently. This phenomenon suggests that synthetic genes which have been designed to include a projected host cell's preferred codons provide a preferred form of foreign genetic material for practice of recombinant DNA techniques. Thus, one aspect of this invention is an adenovirus vector or adenovirus vector in some combination with a vaccine plasmid where both specifically include a gene which is codon optimized for expression in a human cellular environment. As noted herein, a preferred gene for use in the instant invention is a codon-optimized HIV gene and, particularly, HIV gag, pol or nef.

Adenoviral vectors in accordance with the instant invention can be constructed using known techniques, such as those reviewed in Hitt et al, 1997 "Human Adenovirus Vectors for Gene Transfer into Mammalian Cells" Advances in Pharmacology 40:137-206, which is hereby incorporated by reference.

In constructing the adenoviral vectors of this invention, it is often convenient to insert them into a plasmid or shuttle vector. These techniques are known and described in Hitt et al., *supra*. This invention specifically includes both the adenovirus and the adenovirus when inserted into a shuttle plasmid.

Preferred shuttle vectors contain an adenoviral portion and a plasmid portion. The adenoviral portion is essentially the same as the adenovirus vector discussed supra, containing adenoviral sequences (with non-functional or deleted E1 and E3 regions) and the gene expression cassette, flanked by convenient restriction sites. The plasmid portion of the shuttle vector often contains an antibiotic resistance marker under transcriptional control of a prokaryotic promoter so that expression of the antibiotic does not occur in eukaryotic cells. Ampicillin resistance genes, neomycin resistance genes and other pharmaceutically acceptable antibiotic resistance markers may be used. To aid in the high level production of the polynucleotide by

fermentation in prokaryotic organisms, it is advantageous for the shuttle vector to contain a prokaryotic origin of replication and be of high copy number. A number of commercially available prokaryotic cloning vectors provide these benefits. It is desirable to remove non-essential DNA sequences. It is also desirable that the vectors not be able to replicate in eukaryotic cells. This minimizes the risk of integration of polynucleotide vaccine sequences into the recipients' genome. Tissue-specific promoters or enhancers may be used whenever it is desirable to limit expression of the polynucleotide to a particular tissue type.

In one embodiment of this invention, the pre-plasmids (e.g., pMRKAd5pol, pMRKAd5nef and pMRKAd5gag were generated by homologous recombination using the MRKHVE3 (and MRKHVO for the E3- version) backbones and the appropriate shuttle vector, as shown for pMRKAd5pol in Figure 22 and for pMRKAd5nef in Figure 23. The plasmid in linear form is capable of replication after entering the PER.C6® cells and virus is produced. The infected cells and media were harvested after viral replication was complete.

Viral vectors can be propagated in various E1 complementing cell lines, including the known cell lines 293 and PER.C6<sup>®</sup>. Both these cell lines express the adenoviral E1 gene product. PER.C6<sup>®</sup> is described in WO 97/00326 (published January 3, 1997) and issued U.S. Patent No. 6,033,908, both of which are hereby incorporated by reference. It is a primary human retinoblast cell line transduced with an E1 gene segment that complements the production of replication deficient (FG) adenovirus, but is designed to prevent generation of replication competent adenovirus by homologous recombination. Cells of particular interest have been stably transformed with a transgene that encodes the AD5E1A and E1B gene, like PER.C6<sup>®</sup>, from 459 bp to 3510 bp inclusive. 293 cells are described in Graham et al., 1977 J. Gen. Virol 36:59-72, which is hereby incorporated by reference. As stated above, consideration must be given to the adenoviral sequences present in the complementing cell line used. It is important that the sequences not overlap with that present in the vector if the possibility of recombination is to be minimized.

It has been found that vectors generated in accordance with the above description are more effective in inducing an immune response and, thus, constitute very promising vaccine candidates. More particularly, it has been found that first generation adenoviral vectors in accordance with the above description carrying a codon-optimized HIV gag gene, regulated with a strong heterologous promoter can be used as human anti-HIV vaccines, and are capable of inducing immune responses.

Standard techniques of molecular biology for preparing and purifying DNA constructs enable the preparation of the DNA immunogens of this invention.

A vaccine composition comprising an adenoviral vector in accordance with the instant invention may contain physiologically acceptable components, such as buffer, normal saline or phosphate buffered saline, sucrose, other salts and polysorbate. One preferred formulation has: 2.5-10 mM TRIS buffer, preferably about 5 mM TRIS buffer; 25-100 mM NaCl, preferably about 75 mM NaCl; 2.5-10% sucrose, preferably about 5% sucrose; 0.01 -2 mM MgCl<sub>2</sub>; and 0.001%-0.01% polysorbate 80 (plant derived). The pH should range from about 7.0-9.0, preferably about 8.0. One skilled in the art will appreciate that other conventional vaccine excipients may also be used it make the formulation. The preferred formulation contains 5mM TRIS, 75 mM NaCl, 5% sucrose, 1mM MgCl<sub>2</sub>, 0.005% polysorbate 80 at pH 8.0 This has a pH and divalent cation composition which is near the optimum for Ad5 stability and minimizes the potential for adsorption of virus to a glass surface. It does not cause tissue irritation upon intramuscular injection. It is preferably frozen until use.

The amount of adenoviral particles in the vaccine composition to be introduced into a vaccine recipient will depend on the strength of the transcriptional and translational promoters used and on the immunogenicity of the expressed gene product. In general, an immunologically or prophylactically effective dose of  $1x10^7$  to  $1x10^{12}$  particles and preferably about  $1x10^{10}$  to  $1x10^{11}$  particles is administered directly into muscle tissue. Subcutaneous injection, intradermal introduction, impression through the skin, and other modes of administration such as intraperitoneal, intravenous, or inhalation delivery are also contemplated. It is also contemplated that booster vaccinations are to be provided. Following vaccination with HIV adenoviral vector, boosting with a subsequent HIV adenoviral vector and/or plasmid may be desirable. Parenteral administration, such as intravenous, intramuscular, subcutaneous or other means of administration of interleukin-12 protein, concurrently with or subsequent to parenteral introduction of the vaccine compositions of this invention is also advantageous.

The adenoviral vector and/or vaccine plasmids of this invention polynucleotide may be unassociated with any proteins, adjuvants or other agents which impact on the recipients' immune system. In this case, it is desirable for the vector to be in a physiologically acceptable solution, such as, but not limited to, sterile saline or sterile buffered saline. Alternatively, the vector may be associated with an adjuvant known in the art to boost immune responses (i.e., a "biologically effective"

adjuvant), such as a protein or other carrier. Vaccine plasmids of this invention may, for instance, be delivered in saline (e.g., PBS) with or without an adjuvant. Preferred adjuvants are Alum or CRL1005 Block Copolymer. Agents which assist in the cellular uptake of DNA, such as, but not limited to, calcium ions, may also be used to advantage. These agents are generally referred to herein as transfection facilitating reagents and pharmaceutically acceptable carriers. Techniques for coating microprojectiles coated with polynucleotide are known in the art and are also useful in connection with this invention.

5

35

This invention also includes a prime and boost regimen wherein a first adenoviral vector is administered, then a booster dose is given. The booster dose may 10 be repeated at selected time intervals. Alternatively, a preferred inoculation scheme comprises priming with a first adenovirus serotype and then boosting with a second adenovirus serotype. More preferably, the inoculation scheme comprises priming with a first adenovirus serotype and then boosting with a second adenovirus serotype, 15 wherein the first and second adenovirus serotypes are classified within separate subgroups of adenoviruses. The above prime/boost schemes are particularly preferred in those situations where a preexisting immunity is identified to the adenoviral vector of choice. In this type of scheme, the individual or population of individuals is primed with an adenovirus of a serotype other than that to which the preexisting 20 immunity is identified. This enables the first adenovirus to effectuate sufficient expression of the transgene while evading existing immunity to the second adenovirus (the boosting adenovirus) and, further, allows for the subsequent delivery of the transgene via the boosting adenovirus to be more effective. Adenovirus serotype 5 is one example of a virus to which such a scheme might be desirable. In accordance 25 with this invention, therefore, one might decide to prime with a non-group C adenovirus (e.g., Ad12, a group A adenovirus, Ad24, a group D adenovirus, or Ad35, a group B adenovirus) to evade anti-Ad5 immunity and then boost with Ad5, a group C adenovirus. Another preferred embodiment involves administration of a different adenovirus (including non-human adenovirus) vaccine followed by administration of the adenoviral vaccines disclosed. In the alternative, a viral antigen of interest can be first delivered via a viral vaccine other than an adenovirus-based vaccine, and then followed with the adenoviral vaccine disclosed. Alternative viral vaccines include but are not limited to pox virus and venezuelan equine encephilitis virus.

A large body of human and animal data supports the importance of cellular immune responses, especially CTL in controlling (or eliminating) HIV infection. In humans, very high levels of CTL develop following primary infection and correlate

with the control of viremia. Several small groups of individuals have been described who are repeatedly exposed to HIV by remain uninfected; CTL has been noted in several of these cohorts. In the SIV model of HIV infection, CTL similarly develops following primary infection, and it has been demonstrated that addition of anti-CD8 monoclonal antibody abrogated this control of infection and leads to disease progression. This invention uses adenoviral vaccines alone or in combination with plasmid vaccines to induce CTL.

The following non-limiting Examples are presented to better illustrate the invention.

10

15

20

25

30

35

#### **EXAMPLE 1**

Removal of the Intron A Portion of the hCMV Promoter

GMP grade pVIInsHIVgag was used as the starting material to amplify the hCMV promoter. PVIInsHIVgag is a plasmid comprising the CMV immediate-early (IE) promoter and intron A, a full-length codon-optimized HIV gag gene, a bovine growth hormone-derived polyadenylation and transcriptional termination sequence, and a minimal pUC backbone; see Montgomery et al., supra for a description of the plasmid backbone. The amplification was performed with primers suitably positioned to flank the hCMV promoter. A 5' primer was placed upstream of the Msc1 site of the hCMV promoter and a 3' primer (designed to contain the BgIII recognition sequence) was placed 3' of the hCMV promoter. The resulting PCR product (using high fidelity Taq polymerase) which encompassed the entire hCMV promoter (minus intron A) was cloned into TOPO PCR blunt vector and then removed by double digestion with Msc1 and BgIII. This fragment was then cloned back into the original GMP grade pV1JnsHIVgag plasmid from which the original promoter, intron A, and the gag gene were removed following Msc1 and BgIII digestion. This ligation reaction resulted in the construction of a hCMV promoter (minus intron A) + bGHpA expression cassette within the original pV1JnsHIVgag vector backbone. This vector is designated pVIInsCMV(no intron).

The FLgag gene was excised from pV1JnsHIVgag using BgIII digestion and the 1,526 bp gene was gel purified and cloned into pV1JnsCMV(no intron) at the BgIII site. Colonies were screened using Sma1 restriction enzymes to identify clones that carried the Flgag gene in the correct orientation. This plasmid, designated pV1JnsCMV(no intron)-FLgag-bGHpA, was fully sequenced to confirm sequence integrity.

Two additional transgenes were also constructed. The plasmid, pV1JnsCMV(no intron)-FLgag-SPA, is identical to pV1JnsCMV(no intron)-FLgag-bGHpA except that the bovine growth hormone polyadenylation signal has been replaced with a short synthetic polyA signal (SPA) of 50 nucleotides in length. The sequence of the SPA is as shown, with the essential components (poly(A) site, (GT)<sub>n</sub>, and (T)<sub>n</sub>; respectively) underlined:

<u>AATAAA</u>AGATCTTTATTTTCATTAGATCT<u>GTGTG TTGGTTTTTTGTGTG</u> (SEQ ID NO:18).

The plasmid, pV1Jns-mCMV-FLgag-bGHpA, is identical to the pV1JnsCMV(no intron)-FLgag-bGHpA except that the hCMV promoter has been removed and replaced with the murine CMV (mCMV) promoter.

Figure 3 diagrammatically shows the new transgene constructs in comparison with the original transgene.

15

20

25

10

5

#### **EXAMPLE 2**

Gag Expression Assay for Modified Gag Transgenes

Gag Elisa was performed on culture supernatants obtained from transient tissue culture transfection experiments in which the two new hCMV-containing plasmid constructs, pV1JnsCMV(no intron)-FLgag-bGHpA and pV1JnsCMV(no intron)-FLgag-SPA, both devoid of intron A, were compared to pV1JnsHIVgag which, as noted above possesses the intron A as part of the hCMV promoter. Table 2 below shows the *in vitro* gag expression data of the new gag plasmids compared with the GMP grade original plasmid. The results displayed in Table 2 show that both of the new hCMV gag plasmid constructs have expression capacities comparable to the original plasmid construct which contains the intron A portion of the hCMV promoter.

Table 2: In vitro DNA transfection of original and new plasmid HTV-1 gag constructs.

| Plasmid                              | μg gag/10e6 COS cells/5μg DNA/48 hr |
|--------------------------------------|-------------------------------------|
| HIVFL-gagPR9901ª                     | 10.8                                |
| PVIIns-hCMV-FLgag-bGHpAb             | 16.6                                |
| pV1Jns-hCMV-FLgag-SPA <sup>b,c</sup> | 12.0                                |

<sup>&</sup>lt;sup>a</sup> GMP grade pV1Jns-hCMVintronA-FLgag-bGHpA.

10

15

20

#### **EXAMPLE 3**

Rodent (Balb/c) Study for Modified gag Transgenes
A rodent study was performed on the two new plasmid constructs
described above – pV1JnsCMV(no intron)-FLgag-bGHpA and pV1JnsCMV(no
intron)-FLgag-SPA - in order to compare them with the construct described above
possessing the intron A portion of the CMV promoter, pV1JnsHIVgag. Gag antibody
and Elispot responses (described in PCT International Application No.
PCT/US00/18332 (WO 01/02607) filed July 3, 2000, claiming priority to U.S.
Provisional Application Serial No. 60/142,631, filed July 6, 1999 and U.S.
Application Serial No. 60/148,981, filed August 13, 1999, all three applications which
are hereby incorporated by reference) were measured. The results displayed in Table
3 below, show that the new plasmid constructs behaved equivalently to the original
construct in Balb/c mice with respect to their antibody and T-cell responses at both
dosages of plasmid DNA tested, 20 μg and 200 μg.

<sup>5</sup> b New plasmid constructions that have the intron A portion removed from the hCMV promoter.

<sup>&</sup>lt;sup>c</sup> In this construct the bGH terminator has been replaced with the short synthetic polyadenylation signal (SPA)

**EXAMPLE 4** 

Table 3: HIV191: Immunogenicity of V1Jns-gag under different promoter and termination control elements.

| DNA®                | Dose,<br>ug <sup>b</sup> |       | Anti-p24 Titers<br>(3 Wk PD1)° |      |       | SFC/10^6 Cells<br>(4 Wk PD1) <sup>d</sup> |        |  |  |
|---------------------|--------------------------|-------|--------------------------------|------|-------|---|--------|--|--|
| Promoter/terminator |                          | GMT   | +SE                            | -SE  | Media | gag197-205                                | p24    |  |  |
| HIVFL-gagPR9901     | 200                      | 12800 | 4652                           | 3412 | 2(2)  | 129(19)                                   | 30(11) |  |  |
| (GMP grade)         | 20                       | 5572  | 1574                           | 1227 | 0     | 56(9)                                     | 25(6)  |  |  |
| pV1Jns-hCMV-        | 200                      | 11143 | 2831                           | 2257 | 0     | 98(5)                                     | 12(6)  |  |  |
| FL-gag-bGHpA        | 20                       | 7352  | 2808                           | 2032 |       | 73(9)                                     | 11(6)  |  |  |
| pV1Jns-hCMV-        | 200                      | 16890 | 5815                           | 4326 | 1(1)  | 94(4)                                     | 26(7)  |  |  |
| FL-gag-SPA          | 20                       | 5971  | 5361                           | 2825 | 0     | 85(17)                                    | 38(10) |  |  |
| Naïve               | 0                        | 123   | 50                             | 36   | 0     | 0   | 0      |  |  |

<sup>&</sup>lt;sup>B</sup>in PBS

5

15

20

Construction of the Modified Shuttle Vector - "MRKpdelE1 Shuttle"

The modifications to the original Ad5 shuttle vector (pdelE1sp1A; a vector comprising Ad5 sequences from basepairs 1-341 and 3524-5798, with a multiple cloning region between nucleotides 341 and 3524 of Ad5, included the following three manipulations carried out in sequential cloning steps as follows:

- (1) The left ITR region was extended to include the *Pac1* site at the junction between the vector backbone and the adenovirus left ITR sequences. This allow for easier manipulations using the bacterial homologous recombination system.
- 10 (2) The packaging region was extended to include sequences of the wild-type (WT) adenovirus from 342 bp to 450 bp inclusive.
  - (3) The area downstream of pIX was extended 13 nucleotides (i.e., nucleotides 3511-3523 inclusive).

These modifications (Figure 4) effectively reduced the size of the E1 deletion without overlapping with any part of the E1A/E1B gene present in the transformed PER.C6<sup>®</sup> cell line. All manipulations were performed by modifying the Ad shuttle vector pdelE1sp1A.

Once the modifications were made to the shuttle vector, the changes were incorporated into the original Ad5 adenovector backbones (pAdHVO and pAdHVE3) by bacterial homologous recombination using *E. coli* BJ5183 chemically competent cells.

bi.m. Injections into both quads, 50 µL per quad

cn=10;GMT, geometric mean titer; SE, standard. error

dn=5, pooled spleens; mean of triplicate wells and standard, deviation, in parentheses;

#### **EXAMPLE 5**

#### Construction of Modified Adenovector Backbones (E3+ and E3-)

5

10

15

20

25

30

The original adenovectors pAdHVO (comprising all Ad5 sequences except those nucleotides encompassing the E1 and E3 regions ) and pADHVE3 (comprising all Ad5 sequences except those nucleotides encompassing the E1 region), were each reconstructed so that they contained the modifications to the E1 region. This was accomplished by digesting the newly modified shuttle vector (MRKpdelE1 shuttle) with Pac1 and BstZ1101 and isolating the 2,734 bp fragment which corresponds to the adenovirus sequence. This fragment was co-transformed with DNA from either Cla1 linearized pAdHVO (E3- adenovector) or Cla1 linearized pAdHVE3 (E3+adenovector) into E. coli BJ5183 competent cells. At least two colonies from each transformation were selected and grown in Terrific™ broth for 6-8 hours until turbidity was reached. DNA was extracted from each cell pellet and then transformed into E. coli XL1 competent cells. One colony from each transformation was selected and grown for plasmid DNA purification. The plasmid was analyzed by restriction digestions to identify correct clones. The modified adenovectors were designated MRKpAdHVO (E3- plasmid) and MRKpAdHVE3 (E3+ plasmid). Virus from these new adenovectors (MRKHVO and MRKHVE3, respectively) as well as the old version of the adenovectors were generated in the PER.C6® cell lines to accommodate the following series of viral competition experiments. In addition, the multiple cloning site of the original shuttle vector contained ClaI, BamHI, Xho I, EcoRV, HindIII, Sal I, and Bgl II sites. This MCS was replaced with a new MCS containing Not I, Cla I, EcoRV and Asc I sites. This new MCS has been transferred to the MRKpAdHVO and MRKpAdHVE3 pre-plasmids along with the modification made to the packaging region and pIX gene.

#### **EXAMPLE 6**

#### Analysis of the Effect of the Packaging Signal Extension

To study the effects of the modifications made to the E1 deletion region, the viruses obtained from the original backbone (pAdHVE3) and the new backbone (MRKpAdHVE3) were mixed together in equal MOI ratios (1:1 and 5:5) and passaged through several rounds; see Figure 5, Expt.#1. Both of the viruses in the experiment contained the E3 gene intact and did not contain a transgene. The only difference between the two viruses was within the region of the E1 deletion.

Following the coinfection of the viruses at P1 (passage 1), the mixtures were propagated through an additional 4 passages at which time the cells were harvested

and the virus extracted and purified by CsCl banding. The viral DNA was extracted and digested with *Hind*III and the digestion products were then radioactively labeled. For the controls, the respective pre-plasmids (pAdHVE3 ("OLD E3+"); MRKpAdHVE3 ("NEW E3+")) were also digested with *Hind*III (and *Pac*1 to remove the vector backbone) and subsequently labeled with [<sup>33</sup>P]dATP. The radioactively labeled digestion products were subjected to gel electrophoresis and the gel was dried down onto Whatman paper before being exposed to autoradiographic film. Figure 6 clearly shows that the new adenovirus which has the addition made to the packaging signal region has a growth advantage compared with the original adenovirus. In the experiments performed (at either ratio tested), only the digestion bands pertaining to the newly modified virus were present. The diagnostic band of size 3,206 (from the new virus) was clearly present. However, there was no evidence of the diagnostic band of size 2,737 bp expected from the original virus.

15

20

25

30

35

10

5

#### EXAMPLE 7

## Analysis of the Effect of the E3 Gene

The second set of the virus competition study involved mixing equal MOI ratio (1:1) of the newly modified viruses, that obtained from MRKpAdHVO and MRKpAdHVE3 (Figure 5, Expt. #2). In this set, both viruses had the new modifications made to the E1 deletion. The first virus (that from MRKpAdHVO) does not contain an E3 gene. The second virus (that from MRKpAdHVE3) does contain the E3 gene. Neither of the viruses contain a transgene. Following coinfection of the viruses, the mixtures were propagated through an additional 4 passages at which time the cells were harvested and the total virus extracted and purified by CsCl banding. The viral DNA was extracted and digested with HindIII and the digestion products were then radioactively labeled. For the controls, the respective pre-plasmids MRKpAdHVO ("NEW E3-"); MRKpAdHVE3 ("NEW E3+") were also digested with HindIII (and Pac1 to remove the vector backbone) and then labeled with [33P]dATP. The radioactively labeled digestion products were subjected to gel electrophoresis and the gel was dried down onto Whatman paper before being exposed to autoradiographic film. Figure 6 shows the results of the viral DNA analysis of the E3+ virus and E3- virus mixing experiment. The diagnostic band corresponding to the E3+ virus (5,665 bp) was present in greater amount compared with the diagnostic band of 3,010 bp corresponding to the E3- virus. This indicates that the virus that contains the E3 gene is able to amplify more rapidly

compared with the virus that does not contain an E3 gene. This increased amplification capacity has been confirmed by growth studies; see Table 4 below.

#### **EXAMPLE 8**

# Construction of the new shuttle vector containing modified gag transgene – "MRKpdelE1-CMV(no intron)-FLgag-bGHpA"

The modified plasmid pV1JnsCMV(no intron)-FLgag-bGHpA was digested with Msc1 overnight and then digested with Sfi1 for 2 hours at 50°C. The DNA was then treated with Mungbean nuclease for 30 mins at 30°C. The DNA mixture was desalted using the Qiaex II kit and then Klenow treated for 30 mins at 37°C to fully blunt the ends of the transgene fragment. The 2,559 bp transgene fragment was then gel purified. The modified shuttle vector (MRKpdelE1 shuttle) was linearized by digestion with EcoRV, treated with calf intestinal phosphatase and the resulting 6,479 bp fragment was then gel purified. The two purified fragments were then ligated together and several dozen clones were screened to check for insertion of the transgene within the shuttle vector. Diagnostic restriction digestion was performed to identify those clones carrying the transgene in the E1 parallel and E1 anti-parallel orientation. This strategy was followed to clone in the other gag transgenes in the MRKpdelE1 shuttle vector.

20

5

10

15

#### **EXAMPLE 9**

#### Construction of the MRK FG Adenovectors

The shuttle vector containing the HIV-1 gag transgene in the E1 parallel orientation, MRKpdelE1-CMV(no intron)-FLgag-bGHpA, was digested with Pac1. 25 The reaction mixture was digested with BsfZ171. The 5,291 bp fragment was purified by gel extraction. The MRKpAdHVE3 plasmid was digested with Cla1 overnight at 37°C and gel purified. About 100 ng of the 5,290 bp shuttle +transgene fragment and ~100 ng of linearized MRKpAdHVE3 DNA were co-transformed into E. coli BJ5183 chemically competent cells. Several clones were selected and grown in 2 ml 30 Terrific™ broth for 6-8 hours, until turbidity was reached. The total DNA from the cell pellet was purified using Qiagen alkaline lysis and phenol chloroform method. The DNA was precipitated with isopropanol and resuspended in 20 µl dH<sub>2</sub>0. A 2 µl aliquot of this DNA was transformed into E. coli XL-1 competent cells. A single colony from each separate transformation was selected and grown overnight in 3 ml 35 LB +100 μg/ml ampicillin. The DNA was isolated using Qiagen columns. A positive clone was identified by digestion with the restriction enzyme BstEII which cleaves

within the gag gene as well as the plasmid backbone. The pre-plasmid clone is designated MRKpAdHVE3+CMV(no intron)-FLgag-bGHpA and is 37,498 bp in size. This strategy was followed to generate E3- and E3+ versions of each of the other gag transgene constructions in both E1 parallel and E1 anti-parallel versions. Figures 7A, 7B and 7C show the various combinations of adenovectors constructed.

#### **EXAMPLE 10**

#### Plasmid Competition Studies

A series of plasmid competition studies was carried out. Briefly, the screening of the various combinations of new constructs was performed by mixing equal amounts of each of two competing plasmids. In the experiment shown in Figure 8A, plasmids containing the same transgene but in different orientations were mixed together to create a "competition" between the two plasmids. The aim was to look at the effects of transgene orientation. In the experiment shown in Figure 8B, plasmids containing different polyadenylation signals (but in the same orientation) were mixed together in equal amounts. The aim was to assess effects of polyA signals. Following the initial transfection, the virus was passaged through ten rounds and the viral DNA analyzed by radioactive restriction analysis.

Analysis of the viral species from the plasmid mixing experiment (Figure 8A) showed that adenovectors which had the transgene inserted in the E1 parallel orientation amplified better and were able to out-compete the adenovirus which had the transgene inserted in the E1 anti-parallel orientation. Viral DNA analysis of the mixtures at passage 3 and certainly at passage 6, showed a greater ratio of the virus carrying the transgene in the E1 parallel orientation compared with the E1 antiparallel version. By passage 10, the only viral species observed was the adenovector with the transgene in the E1 parallel orientation for both transgenes tested (hCMV(no intron)-FLgag-bGHpA and hCMV(no intron)-FLgag-SPA).

Analysis of the viral species from the plasmid mixing experiment #2 (Figure 8B) at passages 3 and 6 showed that the polyadenylation signals tested (bGHpA and SPA) did not have an effect on the growth of the virus. Even at passage 10 the two viral species in the mixture were still present in equal amounts.

30

20

25

5

#### **EXAMPLE 11**

Virus generation of an enhanced adenoviral construct - "MRK Ad5 HIV-1gag"

The results obtained from the competition study allowed us to make the following conclusions: (1) The packaging signal extension is beneficial; (2) Presence of E3 does enhance viral growth; (3) E1 parallel orientation is recommended; and (4) PolyA signals have no effect on the growth of the adenovirus.

MRK Ad5 HIV-1 gag exhibited the most desirable results. This construct contains the hCMV(no intron)-FLgag-bGHpA transgene inserted into the new E3+ adenovector backbone, MRKpAdHVE3, in the E1 parallel orientation. We have designated this adenovector MRK Ad5 HIV-1 gag. This construct was prepared as outlined below:

The pre-plasmid MRKpAdHVE3+CMV(no intron)-FLgag-bGHpA was digested was Pac1 to release the vector backbone and 3.3 µg was transfected by calcium phosphate method (Amersham Pharmacia Biotech.) in a 6 cm dish containing PER.C6® cells at ~60% confluence. Once CPE was reached (7-10 days), the culture was freeze/thawed three times and the cell debris pelleted. 1 ml of this cell lysate was used to infect into a 6 cm dish containing PER.C6® cells at 80-90% confluence. Once CPE was reached, the culture was freeze/thawed three times and the cell debris pelleted. The cell lysate was then used to infect a 15 cm dish containing PER.C6<sup>®</sup> cells at 80-90% confluence. This infection procedure was continued and expanded at passage 6. The virus was then extracted from the cell pellet by CsCl method. Two bandings were performed (3-gradient CsCl followed by a continuous CsCl gradient). Following the second banding, the virus was dialyzed in A105 buffer. Viral DNA was extracted using pronase treatment followed by phenol chloroform. The viral DNA was then digested with *Hind*III and radioactively labeled with [33P]dATP. Following gel electrophoresis to separate the digestion products the gel was dried down on Whatman paper and then subjected to autoradiography. The digestion products were compared with the digestion products from the pre-plasmid (that had been digested with Pac1/HindIII prior to labeling). The expected sizes were observed, indicating that the virus had been successfully rescued. This strategy was used to rescue virus from each of the various adenovector plasmid constructs prepared.

30

25

5

10

15

# EXAMPLE 12 Stability Analyses

5

10

15

20

35

To determine whether the various adenovector constructs (e.g., MRK Ad5 HIV-1 gag) show genetic stability, the viruses were each passaged continually. The viral DNA was analyzed at passages 3, 6 and 10. Each virus maintained its correct genetic structure. In addition, the stability of the MRK Ad5 HIV-1 gag was analyzed under propagation conditions similar to that performed in large scale production. For this analysis, the transfections of MRK Ad5 HIV-1 gag as well as three other adenoviral vectors were repeated and the virus was purified at P3. The three other adenovectors were as follows: (1) that comprising hCMV(no intron)-Flgag with a bGHpA terminator in an E3- adenovector backbone; (2) that comprising hCMV(no intron)-Flgag with a SPA termination signal in an E3+ adenovector backbone, and that comprising a mCMV-Flgag with a bGHpA terminator in an E3+ adenovector backbone. All of the vectors have the transgene inserted in the E1 parallel orientation. Viral DNA was analyzed by radioactive restriction analysis to confirm that it was correct before being delivered to fermentation cell culture for continued passaging in serum-free media. At P5 each of the four viruses were purified and the viral DNA extracted for analysis by the restriction digestion and radiolabeling procedure. This virus has subsequently been used in a series of studies (in vitro gag expression in COS cells, rodent study and rhesus monkey study) as will be described below. The viruses from P5 are shown in Figure 9.

The passaging under serum-free conditions was continued for the MRKHVE3 (transgene-less, obtained from MRKpAdHVE3 pre-plasmid) and the MRKAd5HIV-1gag (obtained from MRKpAdHVE3+CMV(no intron)-FLgag-bGHpA pre-plasmid) viruses. Figure 10 shows viral DNA analysis by radioactive restriction digestion at passage 11 for MRKHVE3, MRKAd5HIV-1gagE3-, and passage 11 and 12 for MRKAd5HIV-1gag. Aside from the first lane which is the DNA marker lane, the next three lanes are virus from the pre-plasmid controls (controls based on the original virus) - MRKpAdHVE3 (also referred to as "pMRKHVE3"), MRKpAdHVE3+CMV(no intron)-FLgag-bGHpA, and pMRKAd5gag(E3-), respectively. As seen in Figure 10, each of the viral DNA samples show the expected bands with no extraneous bands showing. This signifies that there are no major variant adenovirus species present that can be detected by autoradiography.

Figure 11 shows the results of viral competition study between MRKHVE3 and MRKAd5HIV-1gag. These viruses were mixed together at equal MOI (140 viral

particles each; 280 vp total) at passage 6 and continued to be passaged until P11. Aside from the first lane which is the DNA marker lane, the next two lanes are the pre-plasmid controls obtained from MRKpAdHVE3 and MRKpAdHVE3+CMV(no intron)-FLgag-bGHpA. The next two lanes are the viral DNA from the starting viral material at passage six. The last two lanes are the competition studies performed in duplicate. The data in Figure 11 shows the effect the gag transgene in culture. Growth of a MRKAd5gag virus was compared with growth of a "transgene-less" MRKHVE3. These two viruses were infected at the same MOI (i.e. 140 vp each) at passage 6 and then passaged through to passage 11 and the viral pool was analyzed by radioactive restriction analysis. The data shows that one virus did not out compete the other. Therefore, the gag transgene did not show obvious signs of toxicity to the adenovirus.

10

15

20

25

30

Analysis by *Hind*III digestion shows that each virus specie is present in approximately equal amounts. As above, there does not appear to be signs of any extraneous bands. Figure 12 shows higher passage numbers for MRKAd5HIV-1gag grown under serum-containing conditions. The genome integrity again has been maintained and there is no evidence of rearrangements, even at the highest passage level (P21).

Each of the four vectors shown in Figure 9 were analyzed for amplification capacity. Table 4 below shows the QPA analysis used in the estimation of viral amplification ratios at P4. The determination of the amplification ratio for the original HIV-1 gag construct is based on the clinical lot at P12. It has been shown that amplification rates increases with higher passage number for the original virus. The reason for this observation is due to the emergence of variants which exhibit increased growth rates compared to the intact adenovector. With continued passaging of the original Ad gag vector, the level of variants increases and hence amplification rates increase also.

The MRK Ad5 HIV-1 gag virus has also been continually passaged under process conditions (i.e., serum-free media). Viral DNA extracted from passages 11 and 12 show no evidence of rearrangement.

Table 4: Amplification Ratios Based on AEX and QPA Analysis of Virus Amplification from Passage 3 to Passage 4.

| Ad gag construct       | Amplification Ratio |
|------------------------|---------------------|
| MRKAd5gag              | 470                 |
| HCMV-Flgag-bGHpA [E3-] | 115                 |
| HCMV-Flgag-SPA [E3+]   | 320                 |
| mCMV-FLgag-bGHpA [E3+] | 420                 |
| Original construct *   | 40 - 50             |

.

\* This estimation is based on the clinical lot growth characteristics at Passage 12.

#### **EXAMPLE 13**

10

15

20

25

5

Analytical Evaluation of the enhanced Ad5 Constructs

To study the effects of the transgene and the E3 gene on virus amplification, the enhanced adenoviral vector, MRK Ad5 HIV-1 gag, along with its transgene-less version (MRKpAdHVE3) and its E3- version (MRK Ad5 HIV-1 gag E3-), was studied for several passages under serum-free conditions. Table 5A shows the amplification ratios determined for passages P3 to P8 for MRK Ad5 HIV-1 gag. Within a certain MOI range, it has been determined that the virus output is directly proportional to the virus input. Therefore, the greater the number of virus particles per cell at infection, the greater the virus amount produced. Viral amplification ratios, on the other hand, are inversely proportional to the virus input. The lower the virus input, the greater the amplification ratio.

Table 5B shows the amplification rates of the new E3+ vector backbone MRKpAdHVE3. It has a significantly lower rate of amplification compared with the gag transgene containing version. This may be contributed to the larger size MRK Ad5 HIV-1 gag since it contains the transgene. This inclusion of the transgene brings the size of the adenovirus closer to the size of a wild type Ad5 virus. It is well known that adenoviruses amplify best when they are at close to their wild type genomic size.

Wild type Ad5 is 35,935 bp. The MRKpAdHVE3 is 32, 905 bp in length. The enhanced adenovector MRK Ad5 HIV-1 gag is 35,453bp (See Figure 14 for vector map; see also Figure 15A-X show the complete pre-adenoviral vector sequence, which includes an additional 2,021 bp of the vector backbone).

5

10

Table 5C shows the amplification rates of the new E3- gag containing virus MRK Ad5 HIV-1 gag E3-. Once again, this virus shows lower growth rate than the enhanced adenoviral vector. This may be attributed to the decreased sized of this virus (due to the E3 gene deletion) compared with wild type Ad5. The MRK Ad5 HIV-1 gag E3- virus is 32,810 bp in length. This can be compared with the wild type Ad5 which is 35,935 bp and MRK Ad5 HIV-1 gag which is 35,453 bp in length.

**Table 5A:** Amplification ratios determined by AEX and QPA for **MRKAd5gag** over several continuous passaging in serum free media. Following P5, two replicate samples were taken (rep-1 and rep-2) and analyzed.

# MRKAd5gag rep1

|     |           | nl), Vlability (%) | Harvest Time | Cell Passage | Titer              | Titer       | QPA                        | Ratio   | Amplification      | AEX                  |
|-----|-----------|--------------------|--------------|--------------|--------------------|-------------|----------------------------|---------|--------------------|----------------------|
|     | Infection | Harvest            | hpl          | Number       | 10°° vp/ml culture | 10° vp/csil | 10° TCID <sub>20</sub> /ml | AEX:QPA | Ratto              | Internal Control     |
| P4  | 1.49, 81% | 0.58, 50%          | 44           | 46           | 8.7                | 5.9         | 1.72                       | 50      | 470<br>(MOI = 125) |                      |
| P5  | 1.38, 93% | 0.66, 47%          | 48           | 49           | 6.7                | 4.9         | 1.38                       | 49      | 170                |                      |
| P6  | 1.04, 94% | 0.68, 77%          | 47           | 48           | 5.8                | 5.6         | 1.42                       | 41      | 200                |                      |
| P7  | 1.50, 84% | 0.96, 61%          | 49.5         | 50           | 3.9                | 1.4         | 0.97                       | 40      | 50                 |                      |
| P7  | 1.09, 97% | 0.76, 59%          | 50           | 52           | 5.2                | 4.7         | 1.70                       | 81      | 170                |                      |
| P8  | 1.03, 94% | 0.86, 64%          | 47.5         | 54           | 9.0                | 8.7         | 1.10                       | 82      | 310                |                      |
| P9  | 0.89, 95% | 0.99, 73%          | 47.5         | 56           | 4.4                | 4,9         | 1.03                       | 43      | 175                | 3.12<br>2.84         |
| P10 | 1.09, 91% | 1.06, 66%          | 47,5         | 58           | 3.0                | 2,8         | 1.16                       | 26      | 100                | 2.70<br>2.60         |
| Pii | 1.19, 88% | 0.98, 65%          | 47           | 60           | 3.6                | 3.0         | 1.15                       | 31      | 110                | 2.70<br>2.70         |
| P12 | 0.98, 91% | 0.85, 63%          | 47,5         | 47           | 5.4                | 5.5         | 1.20                       | 45      | 200                | 2.85<br>2.60         |
| P13 | 1.00, 88% | 0.70, 67%          | 49           | 49           | 5.8                | 5.8         | 1.11                       | 52      | 210                | 3.18<br>3.18         |
| P14 | 1.94, 92% | 0.88, 67%          | 46           | 53           | 8.6                | 4.4         |                            |         | 160                | 3.28                 |
| P15 | 0.97, 96% | 0.64, 68%          | 47           | 47           | 6.9                | 7.1         |                            |         | 250                | 3.27<br>3.12<br>2.91 |

**Table 5B:** Amplification ratios determined by AEX and QPA for **MRKHVE3** over several continuous passaging in serum free media. **MRKHVE3** is the new vector backbone which does NOT carry a transgene.

### **MRKHVE3**

|     |           | nl), Vlability (%) | Harvest Time | Cell Passage | Titer                          | Titer       | QPA                        | Ratio   | Amplification      | AEX                  |
|-----|-----------|--------------------|--------------|--------------|--------------------------------|-------------|----------------------------|---------|--------------------|----------------------|
|     | Infection | Harvest            | hpl          | Number       | 10 <sup>10</sup> vp/ml culture | 10° vp/cett | 10° TCID <sub>so</sub> /ml | AEX:QPA | Ratio              | Internal Control     |
| P4  | 1.10, 97% | 1.28, 79%          | 49           | 54           | 4.1                            | 3.8         | 1,70                       | 25      | 300<br>(MOI = 125) |                      |
| P5  | 0.92, 89% | 1.18, 77%          | 47           | . 48         | 4.3                            | 4.7         | 1.24                       | 35      | 170                |                      |
| P6  | 1.55, 85% | 1.25, 76%          | 49.5         | 50           | 1.2                            | 0.8         | 0.56                       | 21      | 30                 |                      |
| P6  | 1.09, 97% | 1.11,81%           | 49           | 52           | 4.0                            | 3.6         | 1.16                       | 34      | 130                |                      |
| P7  | 1.17, 91% | 1.22,91%           | 47,5         | 54           | 3.7                            | 3,2         | 0.50                       | 74      | 110                |                      |
| P8  | 0.98, 88% | 1.41, 83%          | 48           | 56           | 2.1                            | 2,1         | 0.47                       | 45      | 75                 | 3.12<br>2.84         |
| P9  | 1.20, 89% | 1.25, 81%          | 47,5         | 58           | 0.8                            | 0.7         | 0.29                       | 28      | 25                 | 2.70<br>2.60         |
| P10 | 0.99, 82% | 1.55, B5%          | 47           | 60           | 2.3                            | 2.3         | 0.43                       | 53      | 80                 | 2.70<br>2.70         |
| P11 | 1.07, 96% | 1.25, 83%          | 48           | 47           | 2.7                            | 2.5         | 0.41                       | 68      | 90                 | 2.86<br>2.60         |
| P12 | 0.80, 91% | 1.14, 80%          | 49.5         | 49           | 5.9                            | 7.4         | 0.48                       | 123     | 260                | 3.18<br>3.18         |
| P13 | 1.98, 95% | 1.14, 85%          | 45.5         | 53           | 5.8                            | 3.0         |                            |         | 110                | 3.28<br>3.27         |
| P14 | 0.97, 96% | 1.03, 98%          | 45.5         | 47           | 9.4                            | 9.7         |                            |         | 350                | 3.12                 |
| P15 | 0.87, 99% | 0.97, 59%          | 49.5         | 49           | 5.3                            | 6.1         |                            |         | 218                | 2.91<br>2.78<br>2.52 |

Table 5C. Amplification ratios determined by AEX and QPA for MRKAd5gag(E3-) over several continuous passaging in serum free media. This construct is identical to the MRKAd5gag construct except that this version is DELETED of the E3 gene.

5

10

15

## MRKAd5qaq(E3-)

|     |           | ni), Viability (%) | Harvest Time | Cell Passage | Titer<br>10 <sup>19</sup> vp/ml culture | Titer       | QPA                        | Ratio   | Amplification    | AEX              |
|-----|-----------|--------------------|--------------|--------------|---|-------------|----------------------------|---------|------------------|------------------|
|     | Infection | Harvest            | h.p.l.       | Number       |   | 10° vp/cell | 10° TCID <sub>so</sub> /ml | AEX:QPA | Ratto            | Internal Control |
| P4  | 1.62, 77% | 1.12, 62%          | 47.5         | 46           | 2.0                                     | 1.2         | 0.92                       | 20      | 100<br>(MOI=125) |                  |
| P5  | 1.16, 92% | 0.62, 43%          | 49           | 49           | 3.3                                     | 2.9         | 0.99                       | 34      | 100              |                  |
| P6  | 1.71, 86% | 0.20, 10%          | 49           | 50           | 4.7                                     | 2.7         | 1.70                       | 28      | 100              |                  |
| P6  | 1.09, 97% | 0.63, 54%          | 49.5         | 52           | 5.4                                     | 5.0         | 1.76                       | 31      | 180              |                  |
| P7  | 1.17, 91% | 0.98, 72%          | 47.50        | 54           | 7.1                                     | 6.1         | 0.67                       | 106     | 220              |                  |
| P8  | 0.98, 88% | 0.77, 48%          | 48           | 56           | 3.1                                     | 3.2         | 0.66                       | 47      | 115              | 3.12<br>2.84     |
| P9  | 1.20, 89% | 1.03, 72%          | 48           | 58           | 1,8                                     | 1.5         | 0.57                       | 32      | 55               | 2.70<br>2.60     |
| P10 | 0.99, 82% | 0.80, 62%          | 46.5         | 60           | 3.2                                     | 3.2         | 0.68                       | 47      | 115              | 2.70<br>2.70     |
| P11 | 1.07, 96% | 0.98, 70%          | 48.5         | 47           | 5.9                                     | 5.5         | 0.68                       | 87      | 200              | 2.86<br>2.60     |
| P12 | 0.80, 91% | 0.67, 59%          | 50           | 49           | 5.1                                     | 6.4         | 0.72                       | 71      | 230              | 3.18<br>3.18     |
| P13 | 1.96, 95% | 0.91, 59%          | 45.5         | 53           | 7.4                                     | 3.8         |                            |         | 135              | 3.28<br>3.27     |
| P14 | 0.97, 96% | 0.81, 74%          | 48           | 47           | 6.8                                     | 7.0         |                            |         | 250              | 3.12<br>2.91     |
| P15 | 0.87, 99% | 0.84, 56%          | 49           | 49           | 4.8                                     | 5.5         |                            |         | 196              | 2.78<br>2.52     |

#### **EXAMPLE 14**

Gag Expression Analysis of the Novel Constructs

In vitro gag analysis of the MRK Ad5 HIV-1 gag and the original HIV-gag vectors (research and clinical lot) show comparable gag expression. The clinical lot shows only a slightly reduced gag expression level. The most noticeable difference is with the mCMV vector. This vector shows roughly 3 fold lower expression levels compared with the other vectors tested (which all contain hCMV promoters). The mCMV-FLgag with bGHpA assay was performed three times using different propagation and purification lots and it consistently exhibited weaker gag expression.

#### **EXAMPLE 15**

Evaluation of MRK Ad5 HIV-1 gag and Other gag-Containing Adenovectors in Balb/c Mice

Cohorts of 10 balb/c mice were vaccinated intramuscularly with escalating doses of MRK Ad5 HIV-1 gag, and the research and clinical lots of original Ad5HIV-1gag. Serum samples were collected 3 weeks post dose 1 and analyzed by anti-p24 sandwich ELISA.

Anti-p24 titers in mice that received MRK Ad5 HIV-1 gag (107 and 109 vp(viral particle) doses) were comparable (Figure 13) to those of the research lot of Ad5HIV-1 gag, for which much of the early rhesus data were generated on. These titers were also comparable when E3 is deleted (MRKAd5hCMVgagbGHpA(E3-)) or SPA is substituted for bGHpA terminator (MRKAd5 hCMV-gag-SPA (E3+)) or murine CMV promoter is used in place of hCMV (MRKAd5 mCMV-gag-bGHpA (E3+)) in the MRKAd5 backbone.

The results shown in Table 7 indicate that the three other vectors (in addition to the preferred vector, MRK Ad5 HIV-1 gag, are also capable of inducing strong anti-gag antibody responses in mice. Interestingly enough, while the mCMV-FLgag construct containing bGHpA and E3+ in an E1 parallel orientation showed lowest gag expression in the COS cell *in vitro* infection (Table 6) in comparison with the other vectors tested, it generated the greatest anti-gag antibody response this *in vivo* Balb/c study. Table 7 also shows a dose response in anti-gag antibody production in both the research and the clinical lot. As expected, the clinical lot shows reduced anti-gag antibody induction at each dosage level compared to the same dosage used for the research lot.

Table 6: In vitro analysis for gag expression in COS cells by Elisa assay.

20

15

-5

10

| Viral Vectors <sup>a</sup>       | μg gag/4.8x10e5 COS/10e8 parts/48hr |
|----------------------------------|-------------------------------------|
| MRKAd5gag <sup>b</sup>           | 1.40                                |
| Clinical lot Ad5gag <sup>c</sup> | 1.28                                |
| Research lot Ad5gag <sup>d</sup> | 1.32                                |
| MCMVFL-gagbGHpA <sup>e</sup>     | 0.42                                |

<sup>&</sup>lt;sup>a</sup> A<sub>260nm</sub> absorbance readings taken for viral particle determinations.

<sup>&</sup>lt;sup>b</sup> MRKAd5gag was produced in serum free conditions and purified at P5.

<sup>&</sup>lt;sup>c</sup> Clinical lot# Ad5gagFN0001

<sup>25</sup> d Research Ad5FLgag lot# 6399

<sup>°</sup> mCMVFL-gagbGHpA was produced in serum free conditions and purified at P5.

Table 7: mHIV020 Anti-p24 Ab Titers in Balb/c mice (n=10) vaccinated with various Adgag constructs and lots (3 week post dose1).

| Group<br>ID | Vaccine   | Dose<br>(vp) | GMT             | SE upper      | SE lower      |
|-------------|---|--------------|-----------------|---------------|---------------|
| 1 2         | <sup>a</sup> MRKAd5gag                          | 10^7<br>10^9 | 25600<br>409600 | 5877<br>94028 | 4780<br>76473 |
| 3           | hCMV FL-gag bGHpA [E3-] →                       | 10^7         | 7352            | 2077          | 1620          |
| 4           |   | 10^9         | 235253          | 59767         | 47659         |
| 5           | hCMV FL-gag SPA [E3+] →                         | 10^7         | 12800           | 9905          | 236           |
| 6           |   | 10^9         | 310419          | 99181         | 75165         |
| 7 8         | <sup>b</sup> mCMV FL-gag bGHpA [E3+] →          | 10^7         | 44572           | 23504         | 15389         |
|             | "   | 10^9         | 941014          | 239068        | 190636        |
| 9           | <sup>c</sup> hCMV FL-gag bGHpA [ <b>E3-</b> ] ← | 10^7<br>10^9 | 3676<br>117627  | 934<br>17491  | 745<br>15227  |
| 11          | research lot hCMV intronA FL-gag bGHpA [E3-] <- | 10^6         | 528             | 262           | 175           |
| 12          |   | 10^7         | 14703           | 5274          | 3882          |
| 13          |   | 10^8         | 58813           | 14942         | 11915         |
| 14          |   | 10^9         | 204800          | 53232         | 42250         |
| 15          | clinical lot hCMVintronA FL-gag bGHpA [E3-] <-  | 10^6         | 230             | 82            | 61            |
| 16          |   | 10^7         | 4222            | 3405          | 1138          |
| 17          |   | 10^8         | 19401           | 3939          | 3274          |
| 18          |   | 10^9         | 89144           | 25187         | 19639         |
| 19          | Naïve   | none         | 93              | 7             | 6             |

\*2x50 µL i.m. (quad) injections/animal

P.I.s: Youil, Chen, Casimiro

Vaccination: T. Toner, Q. Su

Assav: M. Chen

5

10

#### **EXAMPLE 16**

Comparison of Humoral and Cellular Responses Towards the Original Ad-gag Construct with the New MRK Ad5 HIV-1 gag in Rhesus Monkeys

Cohorts of 3 rhesus monkeys were vaccinated intramuscularly with MRK Ad5 HIV-1 gag or the clinical Ad5gag bulk at two doses,  $10^{11}$  vp and  $10^9$  vp. Immunizations were conducted at week 0, 4, and 25. Serum and PBMC samples were collected at selected time points. The serum sample were assayed for anti-p24 Ab titers (using competitive based assay) and the PBMCs for antigen-specific IFN-gamma secretion following overnight stimulation with gag 20-mer peptide pool (via ELISpot assay).

The results shown in Table 8 indicate comparable responses with respect to the generation of anti-gag antibodies. The frequencies of gag-specific T cells in

<sup>&</sup>lt;sup>a</sup>The structure of MRKAd5gag is: hCMVFL-gagbGHpA [E3+] ightarrow The <u>same lot</u> of MRKAd5gag used in this rodent study was used in the Rhesus monkey study (Tables 7 and 8).

<sup>&</sup>lt;sup>b</sup>The same lot of mCMVFL-gagbGHpA[E3+] used in the in vitro study (Table 6) ws used here.

<sup>&</sup>lt;sup>o</sup>This construct was designed by Volker Sandig. It contains a shorter version of the hCMV promoter than that used in the MRK constructs. The adenovector backbone is identical to the original backbone used in the original Adgag vector. Expression at 10e7 dose from this vector is 7 fold lower then the same dose of the MRKAd5gag and 4 fold lower than the research lot.

peripheral blood assummarized in Table 9 demonstrate a strong cellular immune response generated after a single dose with the new construct MRK Ad5 HIV-1 gag. The responses are also boostable with second dose of the same vector. The vector is also able to induce CD8+ T cell responses (as evident by remaining spot counts after CD4+ depletion of PBMCs) which are responsible for cytotoxic activity.

Table 8 Anti-p24 antibody titers (in mMU/mL) in rhesus macaques immunized with

gag-expressing adenovectors (Protocol HIV203).

| Vaccine                                      | Pre       | Wk4           | Wk8     | Wk 12 | Wk 16 | Wk 20       | Wk 25    | Wk 28  |
|--|-----------|---------------|---------|-------|-------|-------------|----------|--------|
| MR KAd5gag°, 10^11 vp                        | 110       | ****          | 11.50   | WKIZ  | WKIO  | VVK ZU      | ** K Z J | WKZO   |
|  |           | 110           |         | 11500 | 70/0  | 01007       |          | 53.500 |
| 97N010                                       | <10       | 118           | 5528    | 11523 | 7062  | 21997       | ND       | 51593  |
| 97N116                                       | <10       | 62            | 772     | 1447  | 1562  | 2174        | ND       | 20029  |
| 98X007                                       | <10       | _ 66          | 3353_   | 6156  | 6845  | 3719        | ND_      | 24031  |
| MR KAd5gag, 10^9 vp                          |           |               |         |       |       |             |          |        |
| 97N120                                       | <10       | 51            | 204     | 318   | 366   | 482         | ND       | 6550   |
| 97N144                                       | <10       | 18            | 118     | 274   | 706   | 888         | ND       | 7136   |
| 98X008                                       | <10       | 15            | 444     | 386   | 996   | 1072        | ND       | 12851  |
| Ad5gag <sup>b</sup> , Clinical Lot, 10^11 vp | <u> </u>  |               |         |       |       |             |          |        |
| 97X001                                       | <10       | 87            | 2579_   | 4718  | 7174  | 7250        | ND       | 69226  |
| 97N146                                       | <10       | 72            | 3604    | 7380  | 7526  | 18906       | ND       | 60283  |
| 98X009                                       | <10       | 78            | 4183    | 3946  | 3124  | 6956        | ND       | 26226  |
| Ad5gag, Clinical Lot, 10^9 vp                |           |               |         |       |       | <del></del> |          | 72     |
| 97N020                                       | <10       | <10           | 143     | 371   | 390   | 1821        | ND       | 17177  |
| 97X003                                       | <10       | <10           | 39      | 93    | 156   | 596         | ND       | 2053   |
| 98X012                                       | <10       | 81            | 342     | 717   | 956   | 1558        | ND       | 11861  |
| MRKAc5gag (hCMV, bGHpA, E3+)                 |           |               |         |       |       |             |          |        |
| barlginal Actigag vector (hCMV/Intro         | n A bGHp/ | 4, E3-), lot# | F N0001 |       |       | _           | L        |        |
| ND, not determined                           |           |               |         |       | _     |             |          |        |

5

Table 9. Number of gag-specific T cells per million peripheral blood mononuclear cells (PBMCs) in rhesus monkeys immunized with gag-expressing adenovectors. Also included are those frequencies in PBMCs depleted of CD4<sup>+</sup> T cells.

| Grp# | Vaccination           | Mankey ID              |       | Wk                 |       | Wk    |       | l Wk        |         | 6 Wk    |       | 5 Wk       |       | 3 Wk        |
|------|-----------------------|------------------------|-------|--------------------|-------|-------|-------|-------------|---------|---------|-------|------------|-------|-------------|
|      | T=0,4,25 wks          |                        | Media | Gog H <sup>b</sup> | Media | Gog H | Media | Goog H      | Media   | Goog H  | Media | Goog H     | Media | Gog H       |
| 1    | MRKAc5gcg<br>10/11 vp | 97ND10<br>97N010(CD4-) | 6 4   | 89<br>38           | 0     | 395   | 0 3   | 1058<br>993 | 0       | 1174    | 3     | 775<br>76  | 4     | 1074<br>594 |
|      | 10°11 VP              | 97N116                 | Ιī    | 396                | 1     | 609   | ŏ     | 534         | 4       | 395     | ī     | 261        | Ď     | 408         |
|      |                       | 97N116(CD4-)           | 11    | 676                | l     | l     | 0     | 593         | l       | ł       | 0     | 184        | 0     | 666         |
|      |                       | 98X007                 | 10    | 579                | 0     | 1304  | 3     | 2193        | 1       | 2118    | 3     | 1588       | 0     | 2113        |
|      |                       | 98X007(CD4-)           | 20    | 965                |       |       | 0     | 2675        |         |         | 0     | 1656       | 0     | 1278        |
| 2    | MRKAdāgag             | 97N120                 | 5     | 275                | 1     | 249   | 4     | 141         | 4       | 119     | 9     | 206        | 4     | 219         |
|      | 10/9 vp               | 97N120(CD4-)           | 11    | 170                | ١.,   |       | 0     | 85          | ١.      |         | 0     | 75<br>98   | 5     | 219<br>373  |
|      |                       | 97N144                 | 3     | 236                | 6     | 438   |       | 318<br>285  | 3       | 256     | ND    | ND ND      | ő     | 625         |
|      |                       | 97N144(CD4-)           | 6     | 148<br>368         | ۱,    | 1090  | 3     | 891         | 4       | 673     | 3     | 473        | 5     | 735         |
|      |                       | 98X008<br>98X008(CD4-) | 14    | 696                | l '   | 1030  | 0     | 1175        | ٦ ا     | ""      | ١٥    | 391        | 4     | 848         |
|      |                       | 90XUU0(CD4-)           | '4    | 575                |       |       | ١     | ''''        | l       |         |       |            |       |             |
| 3    | Adagag dinkal lat     | 97X001                 | D     | 261                | 1     | 485   | 0     | 817         | 0       | 1220b   | 1     | 894        | 0     | 1858        |
|      | 10^11 vp              | 97X001(CD4-)           | 10    | 283                | i •   |       | 3     | 996         | l _     |         | 0     | 1010       | 0     | 1123        |
|      |                       | 97N146                 | 3     | 150                | 1     | 465   | 0     | 339         | 1       | 1272    | 3     | 1238       | 3     | 1785<br>971 |
|      |                       | 97N146(CD4-)           | 6     | 133                | ١ .   | 339   | 0     | 370         | 0       | 896     | 0     | 654<br>384 | 0     | 1748        |
|      |                       | 98X009                 | 0     | 93<br>73           | 3     | 339   | 3     | 559<br>333  | ٠ ا     | 890     | 6     | 225        | 6     | 644         |
|      |                       | 98X009(CD4-)           | 0     | /3                 |       |       |       | 333         |         |         |       |            |       |             |
| 4    | AdSgaag dinical lat   | 97N020                 | 3     | 30                 | 1     | 101   | 0     | 66<br>15    | 0       | 36      | 0     | 26         | 0     | 41          |
|      | 10/9 vp               | 97ND20(CD4-)<br>97X003 | 10    | 29<br>68           | 5     | 134   | 0     | 18          | ١,      | 38      | 1 4   | 38         | ١٧    | 81          |
|      |                       | 97X003(CD4-)           | 9     | 40                 | ľ     |       | ١ŏ    | 1 6         | Ι '     | ~       | 1 6   | 4          | lŏ    | 19          |
|      |                       | 98XD12                 | 5     | 95                 | 3     | 54    | Ιĭ    | 34          | ١٥      | 18      | Ιŏ    | 20         | Ιĭ    | 121         |
| ı    |                       | 98X012(CD4-)           | l ii  | 70                 | ľ     |       | Ö     | ii          | ]       | .       | Ŏ     | 8          | Ó     | 41          |
| 5    | Noïve                 | 96RD41                 | 6     | 8                  | 1     | 1     | 20    | D<br>14     | 0<br>19 | 0<br>15 | 0     | 0<br>15    | 1 24  | 0           |
|      |                       | 053F                   | 14    | 18                 | 5     | 16 .  | س ا   | 14          | l 'Ÿ    | 15      | "     | '3         | 24    | 1           |

Based on either 4x10/5 or 2x10/5 cells per well (depending on spot density)
ND, not determined

Produce or no peolide control

10

15

20

5 Pool of 20-ca peptides overlapping by 10 caland encompassing the page sequence

The adenovectors described herein and, particularly, MRK Ad5 HIV-1 gag, represent very promising HIV-gag adenovectors with respect to their enhanced growth characteristics in both serum and, more importantly, in serum-free media conditions. In comparison with the current HIV-1 gag adenovector construct, MRK Ad5 HIV-1 gag shows a 5-10 fold increased amplification rate. We have shown that it is genetically stable at passage 21. This construct is able to generate significant cellular immune responses *in vivo* even at a relatively low dose of 10^9 vp. The potency of the MRKAd5gag construct is comparable to, if not better than the original HIV-1gag vector as shown in this rhesus monkey study.

# EXAMPLE 17 CODON OPTIMIZED HIV-1 POL AND CODON OPTIMZED HIV-1 POL MODIFICATIONS

The open reading frames for the various synthetic *pol* genes disclosed herein comprise coding sequences for the reverse transcriptase (or RT which consists of a polymerase and RNase H activity) and integrase (IN). The protein sequence is based

on that of Hxb2r, a clonal isolate of IIIB; this sequence has been shown to be closest to the consensus clade B sequence with only 16 nonidentical residues out of 848 (Korber, et al., 1998, Human retroviruses and AIDS, Los Alamos National Laboratory, Los Alamos, New Mexico). The skilled artisan will understand after review of this specification that any available HIV-1 or HIV-2 strain provides a potential template for the generation of HIV pol DNA vaccine constructs disclosed herein. It is further noted that the protease gene is excluded from the DNA vaccine constructs of the present invention to insure safety from any residual protease activity in spite of mutational inactivation. The design of the gene sequences for both wildtype (wt-pol) and inactivated pol (IA-pol) incorporates the use of human preferred ("humanized") codons for each amino acid residue in the sequence in order to maximize in vivo mammalian expression (Lathe, 1985, J. Mol. Biol. 183:1-12). As can be discerned by inspecting the codon usage in SEQ ID NOs: 1, 3, 5 and 7, the following codon usage for mammalian optimization is preferred: Met (ATG), Gly (GGC), Lys (AAG), Trp (TGG), Ser (TCC), Arg (AGG), Val (GTG), Pro (CCC), Thr (ACC), Glu (GAG); Leu (CTG), His (CAC), Ile (ATC), Asn (AAC), Cys (TGC), Ala (GCC), Gln (CAG), Phe (TTC) and Tyr (TAC). For an additional discussion relating to mammalian (human) codon optimization, see WO 97/31115 (PCT/US97/02294), which, as noted elsewhere in this specification, is hereby incorporated by reference. It is intended that the skilled artisan may use alternative versions of codon optimization or may omit this step when generating HIV pol vaccine constructs within the scope of the present invention. Therefore, the present invention also relates to non-codon optimized versions of DNA molecules and associated recombinant adenoviral HIV vaccines which encode the various wild type and modified forms of the HIV Pol protein disclosed herein. However, codon optimization of these constructs is a preferred embodiment of this invention.

5

10

15

20

25

30

35

A particular embodiment of this portion of the invention comprisies codon optimized nucleotide sequences which encode wt-pol DNA constructs (herein, "wt-pol" or "wt-pol (codon optimized))" wherein DNA sequences encoding the protease (PR) activity are deleted, leaving codon optimized "wild type" sequences which encode RT (reverse transcriptase and RNase H activity) and IN integrase activity. A DNA molecule which encodes this protein is disclosed herein as SEQ ID NO:1, the open reading frame being contained from an initiating Met residue at nucleotides 10-12 to a termination codon from nucleotides 2560-2562. SEQ ID NO:1 is as follows:

AGATCTACCA TEGCCCCCAT CTCCCCCATT GAGACTGTGC CTGTGAAGCT GAAGCCTGGCC

ATGGATGGCC CCAAGGTGAA GCAGTGGCCC CTGACTGAGG AGAAGATCAA GGCCCTGGTG

|    | GAAATCTGCA | CTGAGATGGA | GAAGGAGGC  | AAAATCTCCA | AGATTGGCCC | CGAGAACCCC |
|----|------------|------------|------------|------------|------------|------------|
|    | TACAACACCC | CTGTGTTTGC | CATCAAGAAG | AAGGACTCCA | CCAAGTGGAG | GAAGCTGGTG |
|    | GACTTCAGGG | AGCTGAACAA | GAGGACCCAG | GACTTCTGGG | AGGTGCAGCT | GGGCATCCCC |
|    | CACCCCGCTG | GCCTGAAGAA | GAAGAAGTCT | GTGACTGTGC | TGGATGTGGG | GGATGCCTAC |
| 5  | TTCTCTGTGC | CCCTGGATGA | GGACTTCAGG | AAGTACACTG | CCTTCACCAT | CCCCTCCATC |
|    | AACAATGAGA | CCCCTGGCAT | CAGGTACCAG | TACAATGTGC | TGCCCCAGGG | CTGGAAGGGC |
|    | TCCCCTGCCA | TCTTCCAGTC | CTCCATGACC | AAGATCCTGG | AGCCCTTCAG | GAAGCAGAAC |
|    | CCTGACATTG | TGATCTACCA | GTACATGGAT | GACCTGTATG | TGGGCTCTGA | CCTGGAGATT |
|    | GGGCAGCACA | GGACCAAGAT | TGAGGAGCTG | AGGCAGCACC | TGCTGAGGTG | GGGCCTGACC |
| 10 | ACCCCTGACA | AGAAGCACCA | GAAGGAGCCC | CCCTTCCTGT | GGATGGGCTA | TGAGCTGCAC |
|    | CCCGACAAGT | GGACTGTGCA | GCCCATTGTG | CTGCCTGAGA | AGGACTCCTG | GACTGTGAAT |
|    | GACATCCAGA | AGCTGGTGGG | CAAGCTGAAC | TGGGCCTCCC | AAATCTACCC | TGGCATCAAG |
|    | GTGAGGCAGC | TGTGCAAGCT | GCTGAGGGGC | ACCAAGGCCC | TGACTGAGGT | GATCCCCCTG |
|    | ACTGAGGAGG | CTGAGCTGGA | GCTGGCTGAG | AACAGGGAGA | TCCTGAAGGA | GCCTGTGCAT |
| 15 | GGGGTGTACT | ATGACCCCTC | CAAGGACCTG | ATTGCTGAGA | TCCAGAAGCA | GGGCCAGGGC |
|    | CAGTGGACCT | ACCAAATCTA | CCAGGAGCCC | TTCAAGAACC | TGAAGACTGG | CAAGTATGCC |
|    | AGGATGAGGG | GGGCCCACAC | CAATGATGTG | AAGCAGCTGA | CTGAGGCTGT | GCAGAAGATC |
|    | ACCACTGAGT | CCATTGTGAT | CTGGGGCAAG | ACCCCCAAGT | TCAAGCTGCC | CATCCAGAAG |
|    | GAGACCTGGG | AGACCTGGTG | GACTGAGTAC | TGGCAGGCCA | CCTGGATCCC | TGAGTGGGAG |
| 20 | TTTGTGAACA | CCCCCCCCT  | GGTGAAGCTG | TGGTACCAGC | TGGAGAAGGA | GCCCATTGTG |
|    | GGGGCTGAGA | CCTTCTATGT | GGATGGGGCT | GCCAACAGGG | AGACCAAGCT | GGGCAAGGCT |
|    | GGCTATGTGA | CCAACAGGGG | CAGGCAGAAG | GTGGTGACCC | TGACTGACAC | CACCAACCAG |
|    | AAGACTGAGC | TCCAGGCCAT | CTACCTGGCC | CTCCAGGACT | CTGGCCTGGA | GGTGAACATT |
|    | GTGACTGACT | CCCAGTATGC | CCTGGGCATC | ATCCAGGCCC | AGCCTGATCA | GTCTGAGTCT |
| 25 | GAGCTGGTGA | ACCAGATCAT | TGAGCAGCTG | ATCAAGAAGG | AGAAGGTGTA | CCTGGCCTGG |
|    | GTGCCTGCCC | ACAAGGGCAT | TGGGGGCAAT | GAGCAGGTGG | ACAAGCTGGT | GTCTGCTGGC |
|    | ATCAGGAAGG | TGCTGTTCCT | GGATGGCATT | GACAAGGCCC | AGGATGAGCA | TGAGAAGTAC |
|    | CACTCCAACT | GGAGGGCTAT | GGCCTCTGAC | TTCAACCTGC | CCCCTGTGGT | GGCTAAGGAG |
|    | ATTGTGGCCT | CCTGTGACAA | GTGCCAGCTG | AAGGGGGAGG | CCATGCATGG | GCAGGTGGAC |
| 30 | TGCTCCCCTG | GCATCTGGCA | GCTGGACTGC | ACCCACCTGG | AGGGCAAGGT | GATCCTGGTG |
|    | GCTGTGCATG | TGGCCTCCGG | CTACATTGAG | GCTGAGGTGA | TCCCTGCTGA | GACAGGCCAG |
|    | GAGACTGCCT | ACTTCCTGCT | GAAGCTGGCT | GGCAGGTGGC | CTGTGAAGAC | CATCCACACT |
|    | GACAATGGCT | CCAACTTCAC | TGGGGCCACA | GTGAGGGCTG | CCTGCTGGTG | GGCTGGCATC |
|    | AAGCAGGAGT | TTGGCATCCC | CTACAACCCC | CAGTCCCAGG | GGGTGGTGGA | GTCCATGAAC |
| 35 | AAGGAGCTGA | AGAAGATCAT | TGGGCAGGTG | AGGGACCAGG | CTGAGCACCT | GAAGACAGCT |
|    | GTGCAGATGG | CTGTGTTCAT | CCACAACTTC | AAGAGGAAGG | GGGGCATCGG | GGGCTACTCC |

GCTGGGGAGA GGATTGTGGA CATCATTGCC ACAGACATCC AGACCAAGGA GCTCCAGAAG
CAGATCACCA AGATCCAGAA CTTCAGGGTG TACTACAGGG ACTCCAGGAA CCCCTGTGG
AAGGGCCCTG CCAAGCTGCT GTGGAAGGGG GAGGGGGCTG TGGTGATCCA GGACAACTCT
GACATCAAGG TGGTGCCCAG GAGGAAGGCC AAGATCATCA GGGACTATGG CAAGCAGATG
GCTGGGGATG ACTGTGTGCC CTCCAGGCAG GATGAGGACT AAAGCCCGGG CAGATCT (SEQ
ID NO:1).

5

The open reading frame of the wild type pol construct disclosed as SEQ ID NO:1 contains 850 amino acids, disclosed herein as SEQ ID NO:2, as follows: Met Ala Pro Ile Ser Pro Ile Glu Thr Val Pro Val Lys Leu Lys Pro 10 Gly Met Asp Gly Pro Lys Val Lys Gln Trp Pro Leu Thr Glu Glu Lys Ile Lys Ala Leu Val Glu Ile Cys Thr Glu Met Glu Lys Glu Gly Lys Ile Ser Lys Ile Gly Pro Glu Asn Pro Tyr Asn Thr Pro Val Phe Ala Ile Lys Lys Lys Asp Ser Thr Lys Trp Arg Lys Leu Val Asp Phe Arg Glu Leu Asn Lys Arg Thr Gln Asp Phe Trp Glu Val Gln Leu Gly Ile 15 Pro His Pro Ala Gly Leu Lys Lys Lys Ser Val Thr Val Leu Asp Val Gly Asp Ala Tyr Phe Ser Val Pro Leu Asp Glu Asp Phe Arg Lys Tyr Thr Ala Phe Thr Ile Pro Ser Ile Asn Asn Glu Thr Pro Gly Ile Arg Tyr Gln Tyr Asn Val Leu Pro Gln Gly Trp Lys Gly Ser Pro Ala Ile Phe Gln Ser Ser Met Thr Lys Ile Leu Glu Pro Phe Arg Lys Gln 20 Asn Pro Asp Ile Val Ile Tyr Gln Tyr Met Asp Asp Leu Tyr Val Gly Ser Asp Leu Glu Ile Gly Gln His Arg Thr Lys Ile Glu Glu Leu Arg Gln His Leu Leu Arg Trp Gly Leu Thr Thr Pro Asp Lys Lys His Gln Lys Glu Pro Pro Phe Leu Trp Met Gly Tyr Glu Leu His Pro Asp Lys Trp Thr Val Gln Pro Ile Val Leu Pro Glu Lys Asp Ser Trp Thr Val 25 Asn Asp Ile Gln Lys Leu Val Gly Lys Leu Asn Trp Ala Ser Gln Ile Tyr Pro Gly Ile Lys Val Arg Gln Leu Cys Lys Leu Leu Arg Gly Thr Lys Ala Leu Thr Glu Val Ile Pro Leu Thr Glu Glu Ala Glu Leu Glu Leu Ala Glu Asn Arg Glu Ile Leu Lys Glu Pro Val His Gly Val Tyr Tyr Asp Pro Ser Lys Asp Leu Ile Ala Glu Ile Gln Lys Gln Gly Gln 30 Gly Gln Trp Thr Tyr Gln Ile Tyr Gln Glu Pro Phe Lys Asn Leu Lys Thr Gly Lys Tyr Ala Arg Met Arg Gly Ala His Thr Asn Asp Val Lys Gln Leu Thr Glu Ala Val Gln Lys Ile Thr Thr Glu Ser Ile Val Ile Trp Gly Lys Thr Pro Lys Phe Lys Leu Pro Ile Gln Lys Glu Thr Trp Glu Thr Trp Trp Thr Glu Tyr Trp Gln Ala Thr Trp Ile Pro Glu Trp 35 Glu Phe Val Asn Thr Pro Pro Leu Val Lys Leu Trp Tyr Gln Leu Glu Lys Glu Pro Ile Val Gly Ala Glu Thr Phe Tyr Val Asp Gly Ala Ala

Asn Arg Glu Thr Lys Leu Gly Lys Ala Gly Tyr Val Thr Asn Arg Gly Arg Gln Lys Val Val Thr Leu Thr Asp Thr Thr Asn Gln Lys Thr Glu Leu Gln Ala Ile Tyr Leu Ala Leu Gln Asp Ser Gly Leu Glu Val Asn Ile Val Thr Asp Ser Gln Tyr Ala Leu Gly Ile Ile Gln Ala Gln Pro Asp Gln Ser Glu Ser Glu Leu Val Asn Gln Ile Ile Glu Gln Leu Ile Lys Lys Glu Lys Val Tyr Leu Ala Trp Val Pro Ala His Lys Gly Ile Gly Gly Asn Glu Gln Val Asp Lys Leu Val Ser Ala Gly Ile Arg Lys Val Leu Phe Leu Asp Gly Ile Asp Lys Ala Gln Asp Glu His Glu Lys Tyr His Ser Asn Trp Arg Ala Met Ala Ser Asp Phe Asn Leu Pro Pro Val Val Ala Lys Glu Ile Val Ala Ser Cys Asp Lys Cys Gln Leu Lys 10 Gly Glu Ala Met His Gly Gln Val Asp Cys Ser Pro Gly Ile Trp Gln Leu Asp Cys Thr His Leu Glu Gly Lys Val Ile Leu Val Ala Val His Val Ala Ser Gly Tyr Ile Glu Ala Glu Val Ile Pro Ala Glu Thr Gly Gln Glu Thr Ala Tyr Phe Leu Leu Lys Leu Ala Gly Arg Trp Pro Val Lys Thr Ile His Thr Asp Asn Gly Ser Asn Phe Thr Gly Ala Thr Val 15 Arg Ala Ala Cys Trp Trp Ala Gly Ile Lys Gln Glu Phe Gly Ile Pro Tyr Asn Pro Gln Ser Gln Gly Val Val Glu Ser Met Asn Lys Glu Leu Lys Lys Ile Ile Gly Gln Val Arg Asp Gln Ala Glu His Leu Lys Thr Ala Val Gln Met Ala Val Phe Ile His Asn Phe Lys Arg Lys Gly Gly Ile Gly Gly Tyr Ser Ala Gly Glu Arg Ile Val Asp Ile Ile Ala Thr 20 Asp Ile Gln Thr Lys Glu Leu Gln Lys Gln Ile Thr Lys Ile Gln Asn Phe Arg Val Tyr Tyr Arg Asp Ser Arg Asn Pro Leu Trp Lys Gly Pro Ala Lys Leu Leu Trp Lys Gly Glu Gly Ala Val Val Ile Gln Asp Asn Ser Asp Ile Lys Val Val Pro Arg Arg Lys Ala Lys Ile Ile Arg Asp 25 Tyr Gly Lys Gln Met Ala Gly Asp Asp Cys Val Ala Ser Arg Gln Asp Glu Asp (SEO ID NO:2).

The present invention especially relates to an adenoviral vector vaccine which comprises a codon optimized HIV-1 DNA pol construct wherein, in addition to deletion of the portion of the wild type sequence encoding the protease activity, a combination of active site residue mutations are introduced which are deleterious to HIV-1 pol (RT-RH-IN) activity of the expressed protein. Therefore, the present invention preferably relates to an adenoviral HIV-1 DNA pol-based vaccine wherein the construct is devoid of DNA sequences encoding any PR activity, as well as containing a mutation(s) which at least partially, and preferably substantially, abolishes RT, RNase and/or IN activity. One type of HIV-1 pol mutant which is part and parcel of an adenoviral vector vaccine may include but is not limited to a mutated

30

DNA molecule comprising at least one nucleotide substitution which results in a point mutation which effectively alters an active site within the RT, RNase and/or IN regions of the expressed protein, resulting in at least substantially decreased enzymatic activity for the RT, RNase H and/or IN functions of HIV-1 Pol. In a preferred embodiment of this portion of the invention, a HIV-1 DNA pol construct contains a mutation or mutations within the Pol coding region which effectively abolishes RT, RNase H and IN activity. An especially preferable HIV-1 DNA pol construct in a DNA molecule which contains at least one point mutation which alters the active site of the RT, RNase H and IN domains of Pol, such that each activity is at least substantially abolished. Such a HIV-1 Pol mutant will most likely comprise at least one point mutation in or around each catalytic domain responsible for RT, RNase H and IN activity, respectfully. To this end, an especially preferred HIV-1 DNA pol construct is exemplified herein and contains nine codon substitution mutations which results in an inactivated Pol protein (IA Pol: SEO ID NO:4, Figure 17A-C) which has no PR, RT, RNase or IN activity, wherein three such point mutations reside within each of the RT, RNase and IN catalytic domains. Therefore, an especially preferred exemplification is an adenoviral vaccine which comprises, in an appropriate fashion, a DNA molecule which encodes IA-pol, which contains all nine mutations as shown below in Table 1. An additional preferred amino acid residue for substitution is Asp551, localized within the RNase domain of Pol. Any combination of the mutations disclosed herein may suitable and therefore may be utilized as an IA-Pol-based vaccine of the present invention. While addition and deletion mutations are contemplated and within the scope of the invention, the preferred mutation is a point mutation resulting in a substitution of the wild type amino acid with an alternative amino acid residue.

5

10

15

20

25

| Тя | hl | e | 1 |
|----|----|---|---|

|    |       |            | 14010 1   |                 |
|----|-------|------------|-----------|-----------------|
|    | wt aa | aa residue | mutant aa | enzyme function |
|    | Asp   | 112        | Ala       | RT              |
| 30 | Asp   | 187        | Ala       | RT              |
|    | Asp   | 188        | Ala       | RT              |
|    | Asp . | 445        | . Ala .   | RNase H         |
|    | Glu   | 480        | Ala .     | RNase H         |
|    | Asp   | 500        | Ala       | RNase H         |
|    | Asp   | 626        | Ala       | IN              |
|    | Asp   | 678        | Ala       | IN              |
|    | Glu   | 714        | Ala       | IN              |
|    |       |            |           |                 |

It is preferred that point mutations be incorporated into the IApol mutant adenoviral vaccines of the present invention so as to lessen the possibility of altering epitopes in and around the active site(s) of HIV-1 Pol.

To this end, SEQ ID NO:3 discloses the nucleotide sequence which codes for a codon optimized pol in addition to the nine mutations shown in Table 1, disclosed as follows, and referred to herein as "IApol":

5

10

15

20

25

30

35

AGATCTACCA TGGCCCCCAT CTCCCCCATT GAGACTGTGC CTGTGAAGCT GAAGCCTGGC ATGGATGGCC CCAAGGTGAA GCAGTGGCCC CTGACTGAGG AGAAGATCAA GGCCCTGGTG GAAATCTGCA CTGAGATGGA GAAGGAGGGC AAAATCTCCA AGATTGGCCC CGAGAACCCC TACAACACCC CTGTGTTTGC CATCAAGAAG AAGGACTCCA CCAAGTGGAG GAAGCTGGTG GACTTCAGGG AGCTGAACAA GAGGACCCAG GACTTCTGGG AGGTGCAGCT GGGCATCCCC CACCCCGCTG GCCTGAAGAA GAAGAAGTCT GTGACTGTGC TGGCTGTGGG GGATGCCTAC TTCTCTGTGC CCCTGGATGA GGACTTCAGG AAGTACACTG CCTTCACCAT CCCCTCCATC AACAATGAGA CCCCTGGCAT CAGGTACCAG TACAATGTGC TGCCCCAGGG CTGGAAGGGC TCCCCTGCCA TCTTCCAGTC CTCCATGACC AAGATCCTGG AGCCCTTCAG GAAGCAGAAC CCTGACATTG TGATCTACCA GTACATGGCT GCCCTGTATG TGGGCTCTGA CCTGGAGATT GGGCAGCACA GGACCAAGAT TGAGGAGCTG AGGCAGCACC TGCTGAGGTG GGGCCTGACC ACCCCTGACA AGAAGCACCA GAAGGAGCCC CCCTTCCTGT GGATGGGCTA TGAGCTGCAC CCCGACAAGT GGACTGTGCA GCCCATTGTG CTGCCTGAGA AGGACTCCTG GACTGTGAAT GACATCCAGA AGCTGGTGGG CAAGCTGAAC TGGGCCTCCC AAATCTACCC TGGCATCAAG GTGAGGCAGC TGTGCAAGCT GCTGAGGGGC ACCAAGGCCC TGACTGAGGT GATCCCCCTG ACTGAGGAGG CTGAGCTGGA GCTGGCTGAG AACAGGGAGA TCCTGAAGGA GCCTGTGCAT GGGGTGTACT ATGACCCCTC CAAGGACCTG ATTGCTGAGA TCCAGAAGCA GGGCCAGGGC CAGTGGACCT ACCAAATCTA CCAGGAGCCC TTCAAGAACC TGAAGACTGG CAAGTATGCC AGGATGAGGG GGGCCCACAC CAATGATGTG AAGCAGCTGA CTGAGGCTGT GCAGAAGATC ACCACTGAGT CCATTGTGAT CTGGGGCAAG ACCCCCAAGT TCAAGCTGCC CATCCAGAAG GAGACCTGGG AGACCTGGTG GACTGAGTAC TGGCAGGCCA CCTGGATCCC TGAGTGGGAG TTTGTGAACA CCCCCCCT GGTGAAGCTG TGGTACCAGC TGGAGAAGGA GCCCATTGTG GGGGCTGAGA CCTTCTATGT GGCTGGGGCT GCCAACAGGG AGACCAAGCT GGGCAAGGCT GGCTATGTGA CCAACAGGGG CAGGCAGAAG GTGGTGACCC TGACTGACAC CACCAACCAG AAGACTGCCC TCCAGGCCAT CTACCTGGCC CTCCAGGACT CTGGCCTGGA GGTGAACATT GTGACTGCCT CCCAGTATGC CCTGGGCATC ATCCAGGCCC AGCCTGATCA GTCTGAGTCT GTGCCTGCCC ACAAGGGCAT TGGGGGCAAT GAGCAGGTGG ACAAGCTGGT GTCTGCTGGC ATCAGGAAGG TGCTGTTCCT GGATGGCATT GACAAGGCCC AGGATGAGCA TGAGAAGTAC CACTCCAACT GGAGGGCTAT GGCCTCTGAC TTCAACCTGC CCCCTGTGGT GGCTAAGGAG

```
ATTGTGGCCT CCTGTGACAA GTGCCAGCTG AAGGGGAGG CCATGCATGG GCAGGTGGAC
TGCTCCCCTG GCATCTGGCA GCTGGCCTGC ACCCACCTGG AGGGCAAGGT GATCCTGGTG
GCTGTGCATG TGGCCTCCGG CTACATTGAG GCTGAGGTGA TCCCTGCTGA GACAGGCCAG
GAGACTGCCT ACTTCCTGCT GAAGCTGGCT GGCAGGTGGC CTGTGAAGAC CATCCACACT

5 GCCAATGGCT CCAACTTCAC TGGGGCCACA GTGAGGGCTG CCTGCTGGTG GGCTGCCATC
AAGCAGGAGT TTGGCATCCC CTACAACCCC CAGTCCCAGG GGGTGGTGGC CTCCATGAAC
AAGGAGCTGA AGAAGATCAT TGGGCAGGTG AGGGACCAGG CTGAGCACCT GAAGACAGCT
GTGCAGATGG CTGTGTTCAT CCACAACTTC AAGAGGAAGG GGGGCATCGG GGGCTACTCC
GCTGGGGAGA GGATTGTGGA CATCATTGCC ACAGACATCC AGACCAAGGA GCTCCAGAAG
10 CAGATCACCA AGATCCAGAA CTTCAGGGTG TACTACAGGG ACTCCAGGAA CCCCCTGTGG
AAGGGCCCTG CCAAGCTGCT GTGGAAGGGG GAGGGGCTT TGGTGATCCA GGACAACTCT
GACATCAAGG TGGTGCCCAG GAGGAAGGCC AAGATCATCA GGGACTATGG CAAGCAAGTG
GCTGGGGATG ACTGTGTGC CTCCAGGCAG GATGAGGACT AAAGCCCGGG CAGATCT (SEQ ID
NO:3).
```

15 In order to produce the IA-pol-based adenoviral vaccines of the present invention, inactivation of the enzymatic functions was achieved by replacing a total of nine active site residues from the enzyme subunits with alanine side-chains. As shown in Table 1, all residues that comprise the catalytic triad of the polymerase, namely Asp112, Asp187, and Asp188, were substituted with alanine (Ala) residues 20 (Larder, et al., Nature 1987, 327: 716-717; Larder, et al., 1989, Proc. Natl. Acad. Sci. 1989, 86: 4803-4807). Three additional mutations were introduced at Asp445, Glu480 and Asp500 to abolish RNase H activity (Asp551 was left unchanged in this IA Pol construct), with each residue being substituted for an Ala residue, respectively (Davies, et al., 1991, Science 252:, 88-95; Schatz, et al., 1989, FEBS Lett. 257; 311-25 314; Mizrahi, et al., 1990, Nucl. Acids. Res. 18: pp. 5359-5353). HIV pol integrase function was abolished through three mutations at Asp626, Asp678 and Glu714. Again, each of these residues has been substituted with an Ala residue (Wiskerchen. et al., 1995, J. Virol. 69: 376-386; Leavitt, et al., 1993, J. Biol. Chem. 268: 2113-2119). Amino acid residue Pro3 of SEQ ID NO:4 marks the start of the RT gene. The complete amino acid sequence of IA-Pol is disclosed herein as SEQ ID NO:4 and 30

Met Ala Pro Ile Ser Pro Ile Glu Thr Val Pro Val Lys Leu Lys Pro Gly Met Asp Gly Pro Lys Val Lys Gln Trp Pro Leu Thr Glu Glu Lys Ile Lys Ala Leu Val Glu Ile Cys Thr Glu Met Glu Lys Glu Gly Lys Ile Ser Lys Ile Gly Pro Glu Asn Pro Tyr Asn Thr Pro Val Phe Ala Ile Lys Lys Lys Asp Ser Thr Lys Trp Arg Lys Leu Val Asp Phe Arg

Figure 17A-C, as follows:

.Glu Leu Asn Lys Arg Thr Gln Asp Phe Trp Glu Val Gln Leu Gly Ile Pro His Pro Ala Gly Leu Lys Lys Lys Ser Val Thr Val Leu Ala Val Gly Asp Ala Tyr Phe Ser Val Pro Leu Asp Glu Asp Phe Arg Lys Tyr Thr Ala Phe Thr Ile Pro Ser Ile Asn Asn Glu Thr Pro Gly Ile Arg Tyr Gln Tyr Asn Val Leu Pro Gln Gly Trp Lys Gly Ser Pro Ala Ile Phe Gln Ser Ser Met Thr Lys Ile Leu Glu Pro Phe Arg Lys Gln Asn Pro Asp Ile Val Ile Tyr Gln Tyr Met Ala Ala Leu Tyr Val Gly Ser Asp Leu Glu Ile Gly Gln His Arg Thr Lys Ile Glu Glu Leu Arg Gln His Leu Leu Arg Trp Gly Leu Thr Thr Pro Asp Lys Lys His Gln 10 Lys Glu Pro Pro Phe Leu Trp Met Gly Tyr Glu Leu His Pro Asp Lys Trp Thr Val Gln Pro Ile Val Leu Pro Glu Lys Asp Ser Trp Thr Val Asn Asp Ile Gln Lys Leu Val Gly Lys Leu Asn Trp Ala Ser Gln Ile Tyr Pro Gly Ile Lys Val Arg Gln Leu Cys Lys Leu Leu Arg Gly Thr Lys Ala Leu Thr Glu Val Ile Pro Leu Thr Glu Glu Ala Glu Leu Glu 15 Leu Ala Glu Asn Arg Glu Ile Leu Lys Glu Pro Val His Gly Val Tyr Tyr Asp Pro Ser Lys Asp Leu Ile Ala Glu Ile Gln Lys Gln Gly Gln Gly Gln Trp Thr Tyr Gln Ile Tyr Gln Glu Pro Phe Lys Asn Leu Lys Thr Gly Lys Tyr Ala Arg Met Arg Gly Ala His Thr Asn Asp Val Lys Gln Leu Thr Glu Ala Val Gln Lys Ile Thr Thr Glu Ser Ile Val Ile 20 Trp Gly Lys Thr Pro Lys Phe Lys Leu Pro Ile Gln Lys Glu Thr Trp Glu Thr Trp Trp Thr Glu Tyr Trp Gln Ala Thr Trp Ile Pro Glu Trp Glu Phe Val Asn Thr Pro Pro Leu Val Lys Leu Trp Tyr Gln Leu Glu Lys Glu Pro Ile Val Gly Ala Glu Thr Phe Tyr Val Ala Gly Ala Ala Asn Arg Glu Thr Lys Leu Gly Lys Ala Gly Tyr Val Thr Asn Arg Gly 25 Arg Gln Lys Val Val Thr Leu Thr Asp Thr Thr Asn Gln Lys Thr Ala Leu Gln Ala Ile Tyr Leu Ala Leu Gln Asp Ser Gly Leu Glu Val Asn Ile Val Thr Ala Ser Gln Tyr Ala Leu Gly Ile Ile Gln Ala Gln Pro Asp Gln Ser Glu Ser Glu Leu Val Asn Gln Ile Ile Glu Gln Leu Ile Lys Lys Glu Lys Val Tyr Leu Ala Trp Val Pro Ala His Lys Gly Ile Gly Gly Asn Glu Gln Val Asp Lys Leu Val Ser Ala Gly Ile Arg Lys . Val Leu Phe Leu Asp Gly Ile Asp Lys Ala Gln Asp Glu His Glu Lys Tyr His Ser Asn Trp Arg Ala Met Ala Ser Asp Phe Asn Leu Pro Pro Val Val Ala Lys Glu Ile Val Ala Ser Cys Asp Lys Cys Gln Leu Lys Gly Glu Ala Met His Gly Gln Val Asp Cys Ser Pro Gly Ile Trp Gln 35 Leu Ala Cys Thr His Leu Glu Gly Lys Val Ile Leu Val Ala Val His Val Ala Ser Gly Tyr Ile Glu Ala Glu Val Ile Pro Ala Glu Thr Gly

Gln Glu Thr Ala Tyr Phe Leu Leu Lys Leu Ala Gly Arg Trp Pro Val Lys Thr Ile His Thr Ala Asn Gly Ser Asn Phe Thr Gly Ala Thr Val Arg Ala Ala Cys Trp Trp Ala Gly Ile Lys Gln Glu Phe Gly Ile Pro Tyr Asn Pro Gln Ser Gln Gly Val Val Ala Ser Met Asn Lys Glu Leu Lys Lys Ile Ile Gly Gln Val Arg Asp Gln Ala Glu His Leu Lys Thr Ala Val Gln Met Ala Val Phe Ile His Asn Phe Lys Arg Lys Gly Gly Ile Gly Gly Tyr Ser Ala Gly Glu Arg Ile Val Asp Ile Ile Ala Thr Asp Ile Gln Thr Lys Glu Leu Gln Lys Gln Ile Thr Lys Ile Gln Asn Phe Arg Val Tyr Tyr Arg Asp Ser Arg Asn Pro Leu Trp Lys Gly Pro Ala Lys Leu Leu Trp Lys Gly Glu Gly Ala Val Val Ile Gln Asp Asp Tyr Gly Lys Gln Met Ala Gly Asp Asp Cys Val Ala Ser Arg Gln Asp Glu Asp (SEQ ID NO:4).

As noted above, it will be understood that any combination of the mutations disclosed above may be suitable and therefore be utilized as an IA-pol-based adenoviral HIV vaccine of the present invention, either when administered alone or in a combined modality regime and/or a prime-boost regimen. For example, it may be possible to mutate only 2 of the 3 residues within the respective reverse transcriptase, RNase-H, and integrase coding regions while still abolishing these enzymatic activities. However, the IA-pol construct described above and disclosed as SEQ ID NO:3, as well as the expressed protein (SEQ ID NO:4;) is preferred. It is also preferred that at least one mutation be present in each of the three catalytic domains.

Another aspect of this portion of the invention are codon optimized HIV-1 Pol-based vaccine constructions which comprise a eukaryotic trafficking signal peptide such as from tPA (tissue-type plasminogen activator) or by a leader peptide such as is found in highly expressed mammalian proteins such as immunoglobulin leader peptides. Any functional leader peptide may be tested for efficacy. However, a preferred embodiment of the present invention, as with HIV-1 Nef constructs shown herein, is to provide for a HIV-1 Pol mutant adenoviral vaccine construction wherein the pol coding region or a portion thereof is operatively linked to a leader peptide, preferably a leader peptide from human tPA. In other words, a codon optimized HIV-1 Pol mutant such as IA-Pol (SEQ ID NO:4) may also comprise a leader peptide at the amino terminal portion of the protein, which may effect cellular trafficking and hence, immunogenicity of the expressed protein within the host cell. As noted in Figure 16A-B, a DNA vector which may be utilized to practice the present invention may be modified by known recombinant DNA methodology to contain a leader signal

peptide of interest, such that downstream cloning of the modified HIV-1 protein of interest results in a nucleotide sequence which encodes a modified HIV-1 tPA/Pol protein. In the alternative, as noted above, insertion of a nucleotide sequence which encodes a leader peptide may be inserted into a DNA vector housing the open reading frame for the Pol protein of interest. Regardless of the cloning strategy, the end result is a polynucleotide vaccine which comprises vector components for effective gene expression in conjunction with nucleotide sequences which encode a modified HIV-1 Pol protein of interest, including but not limited to a HIV-1 Pol protein which contains a leader peptide. The amino acid sequence of the human tPA leader utilized herein is as follows: MDAMKRGLCCVLLLCGAVFVSPSEISS (SEQ ID NO:17). Therefore, another aspect of the present invention is to generate HIV-1 Pol-based vaccine constructions which comprise a eukaryotic trafficking signal peptide such as from tPA. To this end, the present invention relates to a DNA molecule which encodes a codon optimized wt-pol DNA construct wherein the protease (PR) activity is deleted and a human tPA leader sequence is fused to the 5' end of the coding region. A DNA molecule which encodes this protein is disclosed herein as SEQ ID NO:5, the open reading frame disclosed herein as SEQ ID NO:6.

10

15

20

25

30

35

To this end, the present invention relates to a DNA molecule which encodes a codon optimized wt-pol DNA construct wherein the protease (PR) activity is deleted and a human tPA leader sequence is fused to the 5' end of the coding region (herein, "tPA-wt-pol"). A DNA molecule which encodes this protein is disclosed herein as SEQ ID NO:5, the open reading frame being contained from an initiating Met residue at nucleotides 8-10 to a termination codon from nucleotides 2633-2635. SEQ ID NO:5 is as follows:

GATCACCATG GATGCAATGA AGAGAGGGCT CTGCTGTGT CTGCTGTGT GTGGAGCAGT
CTTCGTTTCG CCCAGCGAGA TCTCCGCCCC CATCTCCCC ATTGAGACTG TGCCTGTGAA
GCTGAAGCCT GGCATGGATG GCCCCAAGGT GAAGCAGTGG CCCCTGACTG AGGAGAAGAT
CAAGGCCCTG GTGGAAATCT GCACTGAGAT GGAGAAGGAG GGCAAAAATCT CCAAGATTGG
CCCCGAGAAC CCCTACAACA CCCCTGTGTT TGCCATCAAG AAGAAGGACT CCACCAAGTG
GAGGAAGCTG GTGGACTTCA GGGAGCTGAA CAAGAGGACC CAGGACTTCT GGGAGGTGCA
GCTGGGCATC CCCCACCCCG CTGGCCTGAA GAAGAAGAAG TCTGTGACTG TGCTGGATGT
GGGGGATGCC TACTTCTCTG TGCCCCTGGA TGAGGACTTC AGGAAGTACA CTGCCTTCAC
CATCCCCTCC ATCAACAATG AGACCCCTGG CATCAGGTAC CAGTACAATG TGCTGCCCCA
GGGCTGGAAG GGCTCCCCTG CCATCTTCCA GTCCTCCATG ACCAAGATCC TGGAGCCCTT
CAGGAAGCAG AACCCTGACA TTGTGATCTA CCAGTACATG GATGACCTGT ATGTGGGCTC
TGACCTGGAG ATTGGGCAGC ACAGGACCAA GATTGAGGAG CTGAGGCAGC ACCTGCTGAG

GTGGGGCCTG ACCACCCCTG ACAAGAAGCA CCAGAAGGAG CCCCCCTTCC TGTGGATGGG CTATGAGCTG CACCCCGACA AGTGGACTGT GCAGCCCATT GTGCTGCCTG AGAAGGACTC CTGGACTGTG AATGACATCC AGAAGCTGGT GGGCAAGCTG AACTGGGCCT CCCAAATCTA CCCTGGCATC AAGGTGAGGC AGCTGTGCAA GCTGCTGAGG GGCACCAAGG CCCTGACTGA 5 GGTGATCCCC CTGACTGAGG AGGCTGAGCT GGAGCTGGCT GAGAACAGGG AGATCCTGAA GGAGCCTGTG CATGGGGTGT ACTATGACCC CTCCAAGGAC CTGATTGCTG AGATCCAGAA GCAGGGCCAG GGCCAGTGGA CCTACCAAAT CTACCAGGAG CCCTTCAAGA ACCTGAAGAC TGGCAAGTAT GCCAGGATGA GGGGGGCCCA CACCAATGAT GTGAAGCAGC TGACTGAGGC TGTGCAGAAG ATCACCACTG AGTCCATTGT GATCTGGGGC AAGACCCCCA AGTTCAAGCT 10 GCCCATCCAG AAGGAGACCT GGGAGACCTG GTGGACTGAG TACTGGCAGG CCACCTGGAT CCCTGAGTGG GAGTTTGTGA ACACCCCCC CCTGGTGAAG CTGTGGTACC AGCTGGAGAA GGAGCCCATT GTGGGGGCTG AGACCTTCTA TGTGGATGGG GCTGCCAACA GGGAGACCAA GCTGGGCAAG GCTGGCTATG TGACCAACAG GGGCAGGCAG AAGGTGGTGA CCCTGACTGA CACCACCAAC CAGAAGACTG AGCTCCAGGC CATCTACCTG GCCCTCCAGG ACTCTGGCCT 15 GGAGGTGAAC ATTGTGACTG ACTCCCAGTA TGCCCTGGGC ATCATCCAGG CCCAGCCTGA TCAGTCTGAG TCTGAGCTGG TGAACCAGAT CATTGAGCAG CTGATCAAGA AGGAGAAGGT GTACCTGGCC TGGGTGCCTG CCCACAAGGG CATTGGGGGC AATGAGCAGG TGGACAAGCT GGTGTCTGCT GGCATCAGGA AGGTGCTGTT CCTGGATGGC ATTGACAAGG CCCAGGATGA GCATGAGAAG TACCACTCCA ACTGGAGGGC TATGGCCTCT GACTTCAACC TGCCCCCTGT 20 GGTGGCTAAG GAGATTGTGG CCTCCTGTGA CAAGTGCCAG CTGAAGGGGG AGGCCATGCA TGGGCAGGTG GACTGCTCCC CTGGCATCTG GCAGCTGGAC TGCACCCACC TGGAGGGCAA GGTGATCCTG GTGGCTGTGC ATGTGGCCTC CGGCTACATT GAGGCTGAGG TGATCCCTGC TGAGACAGGC CAGGAGACTG CCTACTTCCT GCTGAAGCTG GCTGGCAGGT GGCCTGTGAA GACCATCCAC ACTGACAATG GCTCCAACTT CACTGGGGCC ACAGTGAGGG CTGCCTGCTG 25 GTGGGCTGGC ATCAAGCAGG AGTTTGGCAT CCCCTACAAC CCCCAGTCCC AGGGGGTGGT GGAGTCCATG AACAAGGAGC TGAAGAAGAT CATTGGGCAG GTGAGGGACC AGGCTGAGCA CCTGAAGACA GCTGTGCAGA TGGCTGTGTT CATCCACAAC TTCAAGAGGA AGGGGGGCAT CGGGGGCTAC TCCGCTGGGG AGAGGATTGT GGACATCATT GCCACAGACA TCCAGACCAA GGAGCTCCAG AAGCAGATCA CCAAGATCCA GAACTTCAGG GTGTACTACA GGGACTCCAG 30 GAACCCCCTG TGGAAGGGCC CTGCCAAGCT GCTGTGGAAG GGGGAGGGGG CTGTGGTGAT CCAGGACAAC TCTGACATCA AGGTGGTGCC CAGGAGGAAG GCCAAGATCA TCAGGGACTA TGGCAAGCAG ATGGCTGGGG ATGACTGTGT GGCCTCCAGG CAGGATGAGG ACTAAAGCCC GGGCAGATCT (SEQ ID NO:5).

The open reading frame of the wild type tPA-pol construct disclosed as SEQ ID NO:5 contains 875 amino acids, disclosed herein as SEQ ID NO:6, as follows:

Met Asp Ala Met Lys Arg Gly Leu Cys Cys Val Leu Leu Leu Cys Gly

35

Ala Val Phe Val Ser Pro Ser Glu Ile Ser Ala Pro Ile Ser Pro Ile Glu Thr Val Pro Val Lys Leu Lys Pro Gly Met Asp Gly Pro Lys Val Lys Gln Trp Pro Leu Thr Glu Glu Lys Ile Lys Ala Leu Val Glu Ile Cys Thr Glu Met Glu Lys Glu Gly Lys Ile Ser Lys Ile Gly Pro Glu Asn Pro Tyr Asn Thr Pro Val Phe Ala Ile Lys Lys Lys Asp Ser Thr Lys Trp Arg Lys Leu Val Asp Phe Arg Glu Leu Asn Lys Arg Thr Gln Asp Phe Trp Glu Val Gln Leu Gly Ile Pro His Pro Ala Gly Leu Lys Lys Lys Lys Ser Val Thr Val Leu Asp Val Gly Asp Ala Tyr Phe Ser Val Pro Leu Asp Glu Asp Phe Arg Lys Tyr Thr Ala Phe Thr Ile Pro Ser Ile Asn Asn Glu Thr Pro Gly Ile Arg Tyr Gln Tyr Asn Val Leu 10 Pro Gln Gly Trp Lys Gly Ser Pro Ala Ile Phe Gln Ser Ser Met Thr Lys Ile Leu Glu Pro Phe Arg Lys Gln Asn Pro Asp Ile Val Ile Tyr Gln Tyr Met Asp Asp Leu Tyr Val Gly Ser Asp Leu Glu Ile Gly Gln His Arg Thr Lys Ile Glu Glu Leu Arg Gln His Leu Leu Arg Trp Gly 15 Leu Thr Thr Pro Asp Lys Lys His Gln Lys Glu Pro Pro Phe Leu Trp Met Gly Tyr Glu Leu His Pro Asp Lys Trp Thr Val Gln Pro Ile Val Leu Pro Glu Lys Asp Ser Trp Thr Val Asn Asp Ile Gln Lys Leu Val Gly Lys Leu Asn Trp Ala Ser Gln Ile Tyr Pro Gly Ile Lys Val Arg Gln Leu Cys Lys Leu Leu Arg Gly Thr Lys Ala Leu Thr Glu Val Ile 20 Pro Leu Thr Glu Glu Ala Glu Leu Glu Leu Ala Glu Asn Arg Glu Ile Leu Lys Glu Pro Val His Gly Val Tyr Tyr Asp Pro Ser Lys Asp Leu Ile Ala Glu Ile Gln Lys Gln Gly Gln Gly Gln Trp Thr Tyr Gln Ile Tyr Gln Glu Pro Phe Lys Asn Leu Lys Thr Gly Lys Tyr Ala Arg Met Arg Gly Ala His Thr Asn Asp Val Lys Gln Leu Thr Glu Ala Val Gln 25 Lys Ile Thr Thr Glu Ser Ile Val Ile Trp Gly Lys Thr Pro Lys Phe Lys Leu Pro Ile Gln Lys Glu Thr Trp Glu Thr Trp Trp Thr Glu Tyr Trp Gln Ala Thr Trp Ile Pro Glu Trp Glu Phe Val Asn Thr Pro Pro Leu Val Lys Leu Trp Tyr Gln Leu Glu Lys Glu Pro Ile Val Gly Ala Glu Thr Phe Tyr Val Asp Gly Ala Ala Asn Arg Glu Thr Lys Leu Gly 30 Lys Ala Gly Tyr Val Thr Asn Arg Gly Arg Gln Lys Val Val Thr Leu Thr Asp Thr Thr Asn Gln Lys Thr.Glu Leu Gln Ala Ile Tyr Leu Ala Leu Gln Asp Ser Gly Leu Glu Val Asn Ile Val Thr Asp Ser Gln Tyr Ala Leu Gly Ile Ile Gln Ala Gln Pro Asp Gln Ser Glu Ser Glu Leu Val Asn Gln Ile Ile Glu Gln Leu Ile Lys Lys Glu Lys Val Tyr Leu 35 Ala Trp Val Pro Ala His Lys Gly Ile Gly Gly Asn Glu Gln Val Asp Lys Leu Val Ser Ala Gly Ile Arg Lys Val Leu Phe Leu Asp Gly Ile

Asp Lys Ala Gln Asp Glu His Glu Lys Tyr His Ser Asn Trp Arg Ala Met Ala Ser Asp Phe Asn Leu Pro Pro Val Val Ala Lys Glu Ile Val Ala Ser Cys Asp Lys Cys Gln Leu Lys Gly Glu Ala Met His Gly Gln Val Asp Cys Ser Pro Gly Ile Trp Gln Leu Asp Cys Thr His Leu Glu 5 Gly Lys Val Ile Leu Val Ala Val His Val Ala Ser Gly Tyr Ile Glu Ala Glu Val Ile Pro Ala Glu Thr Gly Gln Glu Thr Ala Tyr Phe Leu Leu Lys Leu Ala Gly Arg Trp Pro Val Lys Thr Ile His Thr Asp Asn Gly Ser Asn Phe Thr Gly Ala Thr Val Arg Ala Ala Cys Trp Trp Ala Gly Ile Lys Gln Glu Phe Gly Ile Pro Tyr Asn Pro Gln Ser Gln Gly 10 Val Val Glu Ser Met Asn Lys Glu Leu Lys Lys Ile Ile Gly Gln Val Arg Asp Gln Ala Glu His Leu Lys Thr Ala Val Gln Met Ala Val Phe Ile His Asn Phe Lys Arg Lys Gly Gly Ile Gly Gly Tyr Ser Ala Gly Glu Arg Ile Val Asp Ile Ile Ala Thr Asp Ile Gln Thr Lys Glu Leu Gln Lys Gln Ile Thr Lys Ile Gln Asn Phe Arg Val Tyr Tyr Arg Asp 15 Ser Arg Asn Pro Leu Trp Lys Gly Pro Ala Lys Leu Leu Trp Lys Gly Glu Gly Ala Val Val Ile Gln Asp Asn Ser Asp Ile Lys Val Val Pro Arg Arg Lys Ala Lys Ile Ile Arg Asp Tyr Gly Lys Gln Met Ala Gly Asp Asp Cys Val Ala Ser Arg Gln Asp Glu Asp (SEQ ID NO:6).

20

25

30

35

The present invention also relates to a codon optimized HIV-1 Pol mutant contained within a recombinant adenoviral vector such as IA-Pol (SEQ ID NO:4) which comprises a leader peptide at the amino terminal portion of the protein, which may effect cellular trafficking and hence, immunogenicity of the expressed protein within the host cell. Any such adenoviral-based HIV-1 DNA pol mutant disclosed in the above paragraphs is suitable for fusion downstream of a leader peptide, such as a leader peptide including but not limited to the human tPA leader sequence. Therefore, any such leader peptide-based HIV-1 pol mutant construct may include but is not limited to a mutated DNA molecule which effectively alters the catalytic activity of the RT, RNase and/or IN region of the expressed protein, resulting in at least substantially decreased enzymatic activity one or more of the RT, RNase H and/or IN functions of HIV-1 Pol. In a preferred embodiment of this portion of the invention, a leader peptide/HIV-1 DNA pol construct contains a mutation or mutations within the Pol coding region which effectively abolishes RT, RNase H and IN activity. An especially preferable HIV-1 DNA pol construct is a DNA molecule which contains at least one point mutation which alters the active site and catalytic activity within the RT, RNase H and IN domains of Pol, such that each activity is at least substantially abolished, and preferably totally abolished. Such a HIV-1 Pol mutant will most likely

comprise at least one point mutation in or around each catalytic domain responsible for RT, RNase H and IN activity, respectfully. An especially preferred embodiment of this portion of the invention relates to a human tPA leader fused to the IA-Pol protein comprising the nine mutations shown in Table 1. The DNA molecule is disclosed herein as SEQ ID NO:7 and the expressed tPA-IA Pol protein comprises a fusion junction as shown in Figure 18. The complete amino acid sequence of the expressed protein is set forth in SEQ ID NO:8. To this end, SEQ ID NO:7 discloses the nucleotide sequence which codes for a human tPA leader fused to the IA Pol protein comprising the nine mutations shown in Table 1 (herein, "tPA-opt-IApol"). The open 10 reading frame begins with the initiating Met (nucleotides 8-10) and terminates with a "TAA" codon at nucleotides 2633-2635. The nucleotide sequence encoding tPA-IAPol is also disclosed as follows: GATCACCATG GATGCAATGA AGAGAGGGCT CTGCTGTGTG CTGCTGCTGT GTGGAGCAGT CTTCGTTTCG CCCAGCGAGA TCTCCGCCC CATCTCCCCC ATTGAGACTG TGCCTGTGAA 15 GCTGAAGCCT GGCATGGATG GCCCCAAGGT GAAGCAGTGG CCCCTGACTG AGGAGAAGAT CAAGGCCCTG GTGGAAATCT GCACTGAGAT GGAGAAGGAG GGCAAAATCT CCAAGATTGG CCCCGAGAAC CCCTACAACA CCCCTGTGTT TGCCATCAAG AAGAAGGACT CCACCAAGTG GAGGAAGCTG GTGGACTTCA GGGAGCTGAA CAAGAGGACC CAGGACTTCT GGGAGGTGCA GCTGGGCATC CCCCACCCG CTGGCCTGAA GAAGAAGAAG TCTGTGACTG TGCTGGCTGT 20 GGGGGATGCC TACTTCTCTG TGCCCCTGGA TGAGGACTTC AGGAAGTACA CTGCCTTCAC CATCCCCTCC ATCAACAATG AGACCCCTGG CATCAGGTAC CAGTACAATG TGCTGCCCCA GGGCTGGAAG GGCTCCCCTG CCATCTTCCA GTCCTCCATG ACCAAGATCC TGGAGCCCTT CAGGAAGCAG AACCCTGACA TTGTGATCTA CCAGTACATG GCTGCCCTGT ATGTGGGCTC TGACCTGGAG ATTGGGCAGC ACAGGACCAA GATTGAGGAG CTGAGGCAGC ACCTGCTGAG 25 GTGGGGCCTG ACCACCCTG ACAAGAAGCA CCAGAAGGAG CCCCCCTTCC TGTGGATGGG CTATGAGCTG CACCCGACA AGTGGACTGT GCAGCCCATT GTGCTGCCTG AGAAGGACTC CTGGACTGTG AATGACATCC AGAAGCTGGT GGGCAAGCTG AACTGGGCCT CCCAAATCTA CCCTGGCATC AAGGTGAGGC AGCTGTGCAA GCTGCTGAGG GGCACCAAGG CCCTGACTGA GGTGATCCCC CTGACTGAGG AGGCTGAGCT GGAGCTGGCT GAGAACAGGG AGATCCTGAA GGAGCCTGTG CATGGGGTGT ACTATGACCC CTCCAAGGAC CTGATTGCTG AGATCCAGAA GCAGGGCCAG GGCCAGTGGA CCTACCAAAT CTACCAGGAG CCCTTCAAGA ACCTGAAGAC TGGCAAGTAT GCCAGGATGA GGGGGGCCCA CACCAATGAT GTGAAGCAGC TGACTGAGGC TGTGCAGAAG ATCACCACTG AGTCCATTGT GATCTGGGGC AAGACCCCCA AGTTCAAGCT GCCCATCCAG AAGGAGACCT GGGAGACCTG GTGGACTGAG TACTGGCAGG CCACCTGGAT 35 CCCTGAGTGG GAGTTTGTGA ACACCCCCC CCTGGTGAAG CTGTGGTACC AGCTGGAGAA GGAGCCCATT GTGGGGGCTG AGACCTTCTA TGTGGCTGGG GCTGCCAACA GGGAGACCAA

GCTGGGCAAG GCTGGCTATG TGACCAACAG GGGCAGGCAG AAGGTGGTGA CCCTGACTGA CACCACCAAC CAGAAGACTG CCCTCCAGGC CATCTACCTG GCCCTCCAGG ACTCTGGCCT GGAGGTGAAC ATTGTGACTG CCTCCCAGTA TGCCCTGGGC ATCATCCAGG CCCAGCCTGA TCAGTCTGAG TCTGAGCTGG TGAACCAGAT CATTGAGCAG CTGATCAAGA AGGAGAAGGT 5 GTACCTGGCC TGGGTGCCTG CCCACAAGGG CATTGGGGGC AATGAGCAGG TGGACAAGCT GGTGTCTGCT GGCATCAGGA AGGTGCTGTT CCTGGATGGC ATTGACAAGG CCCAGGATGA GCATGAGAAG TACCACTCCA ACTGGAGGGC TATGGCCTCT GACTTCAACC TGCCCCCTGT GGTGGCTAAG GAGATTGTGG CCTCCTGTGA CAAGTGCCAG CTGAAGGGGG AGGCCATGCA TGGGCAGGTG GACTGCTCCC CTGGCATCTG GCAGCTGGCC TGCACCCACC TGGAGGGCAA 10 GGTGATCCTG GTGGCTGTGC ATGTGGCCTC CGGCTACATT GAGGCTGAGG TGATCCCTGC TGAGACAGGC CAGGAGACTG CCTACTTCCT GCTGAAGCTG GCTGGCAGGT GGCCTGTGAA GACCATCCAC ACTGCCAATG GCTCCAACTT CACTGGGGCC ACAGTGAGGG CTGCCTGCTG GTGGGCTGGC ATCAAGCAGG AGTTTGGCAT CCCCTACAAC CCCCAGTCCC AGGGGGTGGT GGCCTCCATG AACAAGGAGC TGAAGAAGAT CATTGGGCAG GTGAGGGACC AGGCTGAGCA 15 CCTGAAGACA GCTGTGCAGA TGGCTGTTT CATCCACAAC TTCAAGAGGA AGGGGGGCAT CGGGGGCTAC TCCGCTGGGG AGAGGATTGT GGACATCATT GCCACAGACA TCCAGACCAA GGAGCTCCAG AAGCAGATCA CCAAGATCCA GAACTTCAGG GTGTACTACA GGGACTCCAG GAACCCCTG TGGAAGGGCC CTGCCAAGCT GCTGTGGAAG GGGGAGGGGG CTGTGGTGAT CCAGGACAAC TCTGACATCA AGGTGGTGCC CAGGAGGAAG GCCAAGATCA TCAGGGACTA 20 TGGCAAGCAG ATGGCTGGGG ATGACTGTGT GGCCTCCAGG CAGGATGAGG ACTAAAGCCC GGGCAGATCT (SEQ ID NO:7).

The open reading frame of the tPA-IA-pol construct disclosed as SEQ ID NO:7 contains 875 amino acids, disclosed herein as tPA-IA-Pol and SEQ ID NO:8, as follows:

Met Asp Ala Met Lys Arg Gly Leu Cys Cys Val Leu Leu Leu Cys Gly Ala Val Phe Val Ser Pro Ser Glu Ile Ser Ala Pro Ile Ser Pro Ile Glu Thr Val Pro Val Lys Leu Lys Pro Gly Met Asp Gly Pro Lys Val Lys Gln Trp Pro Leu Thr Glu Glu Lys Ile Lys Ala Leu Val Glu Ile Cys Thr Glu Met Glu Lys Glu Gly Lys Ile Ser Lys Ile Gly Pro Glu Asn Pro Tyr Asn Thr Pro Val Phe Ala Ile Lys Lys Lys Asp Ser Thr Lys Trp Arg Lys Leu Val Asp Phe Arg Glu Leu Asn Lys Arg Thr Gln Asp Phe Trp Glu Val Gln Leu Gly Ile Pro His Pro Ala Gly Leu Lys Lys Lys Lys Ser Val Thr Val Leu Ala Val Gly Asp Ala Tyr Phe Ser Val Pro Leu Asp Glu Asp Phe Arg Lys Tyr Thr Ala Phe Thr Ile Pro Ser Ile Asn Asn Glu Thr Pro Gly Ile Arg Tyr Gln Tyr Asn Val Leu Pro Gln Gly Trp Lys Gly Ser Pro Ala Ile Phe Gln Ser Ser Met Thr

Lys Ile Leu Glu Pro Phe Arg Lys Gln Asn Pro Asp Ile Val Ile Tyr Gln Tyr Met Ala Ala Leu Tyr Val Gly Ser Asp Leu Glu Ile Gly Gln His Arg Thr Lys Ile Glu Glu Leu Arg Gln His Leu Leu Arg Trp Gly Leu Thr Thr Pro Asp Lys Lys His Gln Lys Glu Pro Pro Phe Leu Trp Met Gly Tyr Glu Leu His Pro Asp Lys Trp Thr Val Gln Pro Ile Val Leu Pro Glu Lys Asp Ser Trp Thr Val Asn Asp Ile Gln Lys Leu Val Gly Lys Leu Asn Trp Ala Ser Gln Ile Tyr Pro Gly Ile Lys Val Arg Gln Leu Cys Lys Leu Leu Arg Gly Thr Lys Ala Leu Thr Glu Val Ile Pro Leu Thr Glu Glu Ala Glu Leu Glu Leu Ala Glu Asn Arg Glu Ile Leu Lys Glu Pro Val His Gly Val Tyr Tyr Asp Pro Ser Lys Asp Leu 10 Ile Ala Glu Ile Gln Lys Gln Gly Gln Gly Gln Trp Thr Tyr Gln Ile Tyr Gln Glu Pro Phe Lys Asn Leu Lys Thr Gly Lys Tyr Ala Arg Met Arg Gly Ala His Thr Asn Asp Val Lys Gln Leu Thr Glu Ala Val Gln Lys Ile Thr Thr Glu Ser Ile Val Ile Trp Gly Lys Thr Pro Lys Phe Lys Leu Pro Ile Gln Lys Glu Thr Trp Glu Thr Trp Trp Thr Glu Tyr 15 Trp Gln Ala Thr Trp Ile Pro Glu Trp Glu Phe Val Asn Thr Pro Pro Leu Val Lys Leu Trp Tyr Gln Leu Glu Lys Glu Pro Ile Val Gly Ala Glu Thr Phe Tyr Val Ala Gly Ala Ala Asn Arg Glu Thr Lys Leu Gly Lys Ala Gly Tyr Val Thr Asn Arg Gly Arg Gln Lys Val Val Thr Leu Thr Asp Thr Thr Asn Gln Lys Thr Ala Leu Gln Ala Ile Tyr Leu Ala 20 Leu Gln Asp Ser Gly Leu Glu Val Asn Ile Val Thr Ala Ser Gln Tyr Ala Leu Gly Ile Ile Gln Ala Gln Pro Asp Gln Ser Glu Ser Glu Leu Val Asn Gln Ile Ile Glu Gln Leu Ile Lys Lys Glu Lys Val Tyr Leu Ala Trp Val Pro Ala His Lys Gly Ile Gly Gly Asn Glu Gln Val Asp Lys Leu Val Ser Ala Gly Ile Arg Lys Val Leu Phe Leu Asp Gly Ile 25 Asp Lys Ala Gln Asp Glu His Glu Lys Tyr His Ser Asn Trp Arg Ala Met Ala Ser Asp Phe Asn Leu Pro Pro Val Val Ala Lys Glu Ile Val Ala Ser Cys Asp Lys Cys Gln Leu Lys Gly Glu Ala Met His Gly Gln Val Asp Cys Ser Pro Gly Ile Trp Gln Leu Ala Cys Thr His Leu Glu Gly Lys Val Ile Leu Val Ala Val His Val Ala Ser Gly Tyr Ile Glu Ala Glu Val Ile Pro Ala Glu Thr Gly Gln Glu Thr Ala Tyr Phe Leu Leu Lys Leu Ala Gly Arg Trp Pro Val Lys Thr Ile His Thr Ala Asn Gly Ser Asn Phe Thr Gly Ala Thr Val Arg Ala Ala Cys Trp Trp Ala Gly Ile Lys Gln Glu Phe Gly Ile Pro Tyr Asn Pro Gln Ser Gln Gly Val Val Ala Ser Met Asn Lys Glu Leu Lys Lys Ile Ile Gly Gln Val Arg Asp Gln Ala Glu His Leu Lys Thr Ala Val Gln Met Ala Val Phe

Ile His Asn Phe Lys Arg Lys Gly Gly Ile Gly Gly Tyr Ser Ala Gly Glu Arg Ile Val Asp Ile Ile Ala Thr Asp Ile Gln Thr Lys Glu Leu Gln Lys Gln Ile Thr Lys Ile Gln Asn Phe Arg Val Tyr Tyr Arg Asp Ser Arg Asn Pro Leu Trp Lys Gly Pro Ala Lys Leu Leu Trp Lys Gly Glu Gly Ala Val Val Ile Gln Asp Asp Ser Asp Ile Lys Val Val Pro Arg Arg Lys Ala Lys Ile Ile Arg Asp Tyr Gly Lys Gln Met Ala Gly Asp Asp Cys Val Ala Ser Arg Gln Asp Glu Asp (SEQ ID NO:8).

### **EXAMPLE 18**

# CODON OPTIMIZED HIV-1 NEF AND CODON OPTIMIZED HIV-1 NEF MODIFICATIONS

10

Codon optimized version of HIV-1 Nef and HIV-1 Nef modifications are essentially as described in U.S. Application Serial No. 09/738,782, filed December 15, 2000 and PCT International Application PCT/US00/34162, also filed 15 December 15, 2000, both documents which are hereby incorporated by reference. As disclosed within the above-mentioned documents, particular embodiments of codon optimized Nef and Nef modifications relate to a DNA molecule encoding HIV-1 Nef from the HIV-1 jfrl isolate wherein the codons are optimized for expression in a mammalian system such as a human. The DNA molecule which encodes this protein 20 is disclosed herein as SEQ ID NO:9, while the expressed open reading frame is disclosed herein as SEQ ID NO:10. Another embodiment of Nef-based coding regions for use in the adenoviral vectors of the present invention comprise a codon optimized DNA molecule encoding a protein containing the human plasminogen activator (tpa) leader peptide fused with the NH<sub>2</sub>-terminus of the HTV-1 Nef polypeptide. The DNA molecule which encodes this protein is disclosed herein as 25 SEQ ID NO:11, while the expressed open reading frame is disclosed herein as SEO ID NO:12. Another modified Nef optimized coding region relates to a DNA molecule encoding optimized HIV-1 Nef wherein the open reading frame codes for modifications at the amino terminal myristylation site (Gly-2 to Ala-2) and 30 substitution of the Leu-174-Leu-175 dileucine motif to Ala-174-Ala-175, herein described as opt nef (G2A, LLAA). The DNA molecule which encodes this protein is disclosed herein as SEQ ID NO:13, while the expressed open reading frame is disclosed herein as SEQ ID NO:14. An additional embodiment relates to a DNA molecule encoding optimized HIV-1 Nef wherein the amino terminal myristylation 35 site and dileucine motif have been deleted, as well as comprising a tPA leader peptide. This DNA molecule, opt tpanef (LLAA), comprises an open reading frame which

encodes a Nef protein containing a tPA leader sequence fused to amino acid residue 6-216 of HIV-1 Nef (jfrl), wherein Leu-174 and Leu-175 are substituted with Ala-174 and Ala-175, herein referred to as opt tpanef (LLAA) is disclosed herein as SEQ ID NO:15, while the expressed open reading frame is disclosed herein as SEQ ID NO:16.

As disclosed in the above-identified documents (U.S. Application Serial No. 09/738,782 and PCT International Application PCT/US00/34162) and reiterated herein, the following nef-based nucleotide and amino acid sequences which comprise the respective open reading frame are as follows:

5

10

15

20

25

30

35

1. The nucleotide sequence of the codon optimized version of HIV-1 jrfl nef gene is disclosed herein as SEQ ID NO:9, as shown herein:

GATCTGCCAC CATGGGCGGC AAGTGGTCCA AGAGGTCCGT GCCCGGCTGG TCCACCGTGA
GGGAGAGGAT GAGGAGGGCC GAGCCCGCCG CCGACAGGGT GAGGAGGACC GAGCCCGCCG
CCGTGGGCGT GGGCGCCGTG TCCAGGGACC TGGAGAAGCA CGGCGCCATC ACCTCCTCCA
ACACCGCCGC CACCAACGCC GACTGCGCCT GGCTGGAGGC CCAGGAGGAC GAGGAGGTGG
GCTTCCCCGT GAGGCCCCAG GTGCCCCTGA GGCCCATGAC CTACAAGGGC GCCGTGGACC
TGTCCCACTT CCTGAAGGAG AAGGGCGGCC TGGAGGGCCT GATCCACTCC CAGAAGAGGC
AGGACATCCT GGACCTGTGG GTGTACCACA CCCAGGGCTA CTTCCCCGAC TGGCAGAACT
ACACCCCCGG CCCCGGCATC AGGTTCCCCC TGACCTTCGG CTGGTGCTTC AAGCTGGTGC
CCGTGGAGCC CGAGAAGGTG GAGGAGGCCA ACGAGGGCGA GAACAACTGC CTGCTGCACC
CCATGTCCCA GCACGGCATC GAGGACCCCG AGAAGGAGGT GCTGGAGTGG AGGTTCGACT
CCAAGCTGGC CTTCCACCAC GTGGCCAGGG AGCTGCACCC CGAGTACTAC AAGGACTGCT
AAAGCCCGGG C (SEQ ID NO:9).

Preferred codon usage is as follows: Met (ATG), Gly (GGC), Lys (AAG), Trp (TGG), Ser (TCC), Arg (AGG), Val (GTG), Pro (CCC), Thr (ACC), Glu (GAG); Leu (CTG), His (CAC), Ile (ATC), Asn (AAC), Cys (TGC), Ala (GCC), Gln (CAG), Phe (TTC) and Tyr (TAC). For an additional discussion relating to mammalian (human) codon optimization, see WO 97/31115 (PCT/US97/02294), which is hereby incorporated by reference. See also Figure 19A-B for a comparion of wild type vs. codon optimized nucleotides comprising the open reading frame of HIV-Nef.

The open reading frame for SEQ ID NO:9 above comprises an initiating methionine residue at nucleotides 12-14 and a "TAA" stop codon from nucleotides 660-662. The open reading frame of SEQ ID NO:9 provides for a 216 amino acid HIV-1 Nef protein expressed through utilization of a codon optimized DNA vaccine vector. The 216 amino acid HIV-1 Nef (jfrl) protein is disclosed herein as SEQ ID NO:10, and as follows:

Met Gly Gly Lys Trp Ser Lys Arg Ser Val Pro Gly Trp Ser Thr Val

Arg Glu Arg Met Arg Arg Ala Glu Pro Ala Ala Asp Arg Val Arg Arg Thr Glu Pro Ala Ala Val Gly Val Gly Ala Val Ser Arg Asp Leu Glu Lys His Gly Ala Ile Thr Ser Ser Asn Thr Ala Ala Thr Asn Ala Asp Cys Ala Trp Leu Glu Ala Gln Glu Asp Glu Glu Val Gly Phe Pro Val Arg Pro Gln Val Pro Leu Arg Pro Met Thr Tyr Lys Gly Ala Val Asp Leu Ile His Ser Gln Lys Arg Gln Asp Ile Leu Asp Leu Trp Val Tyr His Thr Gln Gly Tyr Phe Pro Asp Trp Gln Asn Tyr Thr Pro Gly Pro Gly Ile Arg Phe Pro Leu Thr Pro Leu Thr Pro Gly Pro Glu Pro Glu Lys Phe Glu Lys Phe Glu Lys Phe Lys Leu Val Pro Val Glu Pro Glu Lys Val Glu Glu Ala Asn Glu Gly Glu Asn Asn Cys Leu His Pro Met Ser Gln His Gly Ile Glu Asp Pro Glu Lys Glu Val Leu Glu Glu Trp Arg Phe Asp Ser Lys Leu Ala Phe His His Val Ala Arg Glu Leu His Pro Glu Tyr Tyr Tyr Lys Asp Cys (SEQ ID NO:10).

5

10

15

20

25

30

35

HIV-1 Nef is a 216 amino acid cytosolic protein which associates with the inner surface of the host cell plasma membrane through myristylation of Gly-2 (Franchini et al., 1986, Virology 155: 593-599). While not all possible Nef functions have been elucidated, it has become clear that correct trafficking of Nef to the inner plasma membrane promotes viral replication by altering the host intracellular environment to facilitate the early phase of the HIV-1 life cycle and by increasing the infectivity of progeny viral particles. In one aspect of the invention regarding codon-optimized, protein-modified polypeptides, the nef-encoding region of the adenovirus vector of the present invention is modified to contain a nucleotide sequence which encodes a heterologous leader peptide such that the amino terminal region of the expressed protein will contain the leader peptide. The diversity of function that typifies eukaryotic cells depends upon the structural differentiation of their membrane boundaries. To generate and maintain these structures, proteins must be transported from their site of synthesis in the endoplasmic reticulum to predetermined destinations throughout the cell. This requires that the trafficking proteins display sorting signals that are recognized by the molecular machinery responsible for route selection located at the access points to the main trafficking pathways. Sorting decisions for most proteins need to be made only once as they traverse their biosynthetic pathways since their final destination, the cellular location at which they perform their function, becomes their permanent residence. Maintenance of intracellular integrity depends in part on the selective sorting and accurate transport of proteins to their correct destinations. Defined sequence motifs exist in proteins which can act as 'address labels'. A number of sorting signals have

been found associated with the cytoplasmic domains of membrane proteins. An effective induction of CTL responses often required sustained, high level endogenous expression of an antigen. As membrane-association via myristylation is an essential requirement for most of Nef's function, mutants lacking myristylation, by glycine-to-alanine change, change of the dileucine motif and/or by substitution with a tpa leader sequence as described herein, will be functionally defective, and therefore will have improved safety profile compared to wild-type Nef for use as an HIV-1 vaccine component.

10

15

25

30

35

In another embodiment of this portion of the invention, either the DNA vector or the HIV-1 nef nucleotide sequence is modified to include the human tissue-specific plasminogen activator (tPA) leader. As shown in Figure 16A-B, a DNA vector may be modified by known recombinant DNA methodology to contain a leader signal peptide of interest, such that downstream cloning of the modified HIV-1 protein of interest results in a nucleotide sequence which encodes a modified HIV-1 tPA/Nef protein. In the alternative, as noted above, insertion of a nucleotide sequence which encodes a leader peptide may be inserted into a DNA vector housing the open reading frame for the Nef protein of interest. Regardless of the cloning strategy, the end result is a polynucleotide vaccine which comprises vector components for effective gene expression in conjunction with nucleotide sequences which encode a modified HIV-1 Nef protein of interest, including but not limited to a HIV-1 Nef protein which contains a leader peptide. The amino acid sequence of the human tPA leader utilized herein is as follows: MDAMKRGLCCVLLLCGAVFVSPSEISS (SEQ ID NO:17).

It has been shown that myristylation of Gly-2 in conjunction with a dileucine motif in the carboxy region of the protein is essential for Nef-induced down regulation of CD4 (Aiken et al., 1994, Cell 76: 853-864) via endocytosis. It has also been shown that Nef expression promotes down regulation of MHCI (Schwartz et al., 1996, Nature Medicine 2(3): 338-342) via endocytosis. The present invention relates in part to DNA vaccines which encode modified Nef proteins altered in trafficking and/or functional properties. The modifications introduced into the adenoviral vector HIV vaccines of the present invention include but are not limited to additions, deletions or substitutions to the nef open reading frame which results in the expression of a modified Nef protein which includes an amino terminal leader peptide, modification or deletion of the amino terminal myristylation site, and modification or deletion of the dileucine motif within the Nef protein and which alter function within the infected host cell. Therefore, a central theme of the DNA molecules and recombinant adenoviral HIV vaccines of the present invention is (1)

host administration and intracellular delivery of a codon optimized nef-based adenoviral HIV vaccine; (2) expression of a modified Nef protein which is immunogenic in terms of eliciting both CTL and Th responses; and, (3) inhibiting or at least altering known early viral functions of Nef which have been shown to promote HIV-1 replication and load within an infected host. Therefore, the nef coding region may be altered, resulting in a DNA vaccine which expresses a modified Nef protein wherein the amino terminal Gly-2 myristylation residue is either deleted or modified to express alternate amino acid residues. Also, the nef coding region may be altered so as to result in a DNA vaccine which expresses a modified Nef protein wherein the dileucine motif is either deleted or modified to express alternate amino acid residues. In addition, the adenoviral vector HIV vaccines of the present invention also relate to an isolated DNA molecule, regardless of codon usage, which expresses a wild type or modified Nef protein as described herein, including but not limited to modified Nef proteins which comprise a deletion or substitution of Gly 2, a deletion or substitution of Leu 174 and Leu 175 and/or inclusion of a leader sequence.

Therefore, specific Nef-based constructs further include the following, as exemplification's and not limitations. For example, the present invention relates to an adenoviral vector vaccine which encodes modified forms of HIV-1, an open reading frame which encodes a Nef protein which comprises a tPA leader sequence fused to amino acid residue 6-216 of HIV-1 Nef (jfrl) is referred to herein as opt tpanef. The nucleotide sequence comprising the open reading frame of opt tpanef is disclosed herein as SEQ ID NO:11, as shown below:

CATGGATGCA ATGAAGAGAG GGCTCTGCTG TGTGCTGCTG CTGTGTGGAG CAGTCTTCGT
TTCGCCCAGC GAGATCTCCT CCAAGAGGTC CGTGCCCGGC TGGTCCACCG TGAGGGAGAG
GATGAGGAGG GCCGAGCCCG CCGCCGACAG GGTGAGGAGG ACCGAGCCCG CCGCCGTGGG
CGTGGGCGCC GTGTCCAGGG ACCTGGAGAA GCACGGCGCC ATCACCTCCT CCAACACCGC
CGCCACCAAC GCCGACTGCG CCTGGCTGGA GGCCCAGGAG GACGAGGAGG TGGGCTTCCC
CGTGAGGCCC CAGGTGCCCC TGAGGCCCAT GACCTACAAG GGCGCCGTGG ACCTGTCCCA
CTTCCTGAAG GAGAAGGGCG GCCTGGAGGG CCTGATCCAC TCCCAGAAGA GGCAGGACAT
CCTGGACCTG TGGGTGTACC ACACCCAGGG CTACTTCCCC GACTGGCAGA ACTACACCCC
CGGCCCCGGC ATCAGGTTCC CCCTGACCTT CGGCTGGTGC TTCAAGCTGG TGCCCGTGGA
GCCCGAGAAG GTGGAGGAGG CCAACGAGGG CGAGAACAAC TGCCTGCTGC ACCCCATGTC
CCAGCACGGC ATCGAGGACC CCGAGAAGGA GGTGCTGGAG TGGAGGTTCG ACTCCAAGCT
GGCCTTCCAC CACGTGGCCA GGGAGCTGCA CCCCGAGTAC TACAAGGACT GCTAAAGCC
(SEQ ID NO:11).

The open reading frame for SEQ ID NO:11 comprises an initiating methionine

residue at nucleotides 2-4 and a "TAA" stop codon from nucleotides 713-715. The open reading frame of SEQ ID NO:3 provides for a 237 amino acid HIV-1 Nef protein which comprises a tPA leader sequence fused to amino acids 6-216 of HIV-1 Nef, including the dileucine motif at amino acid residues 174 and 175. This 237 amino acid tPA/Nef (jfrl) fusion protein is disclosed herein as SEQ ID NO:12, and is shown as follows:

Met Asp Ala Met Lys Arg Gly Leu Cys Cys Val Leu Leu Leu Cys Gly Ala Val Phe Val Ser Pro Ser Glu Ile Ser Ser Lys Arg Ser Val Pro Gly Trp Ser Thr Val Arg Glu Arg Met Arg Arg Ala Glu Pro Ala Ala Asp Arg Val Arg Arg Thr Glu Pro Ala Ala Val Gly Val Gly Ala Val 10 Ser Arg Asp Leu Glu Lys His Gly Ala Ile Thr Ser Ser Asn Thr Ala Ala Thr Asn Ala Asp Cys Ala Trp Leu Glu Ala Gln Glu Asp Glu Glu Val Gly Phe Pro Val Arg Pro Gln Val Pro Leu Arg Pro Met Thr Tyr Lys Gly Ala Val Asp Leu Ser His Phe Leu Lys Glu Lys Gly Gly Leu Glu Gly Leu Ile His Ser Gln Lys Arg Gln Asp Ile Leu Asp Leu Trp 15 Val Tyr His Thr Gln Gly Tyr Phe Pro Asp Trp Gln Asn Tyr Thr Pro Gly Pro Gly Ile Arg Phe Pro Leu Thr Phe Gly Trp Cys Phe Lys Leu Val Pro Val Glu Pro Glu Lys Val Glu Glu Ala Asn Glu Gly Glu Asn Asn Cys Leu Leu His Pro Met Ser Gln His Gly Ile Glu Asp Pro Glu Lys Glu Val Leu Glu Trp Arg Phe Asp Ser Lys Leu Ala Phe His His 20 Val Ala Arg Glu Leu His Pro Glu Tyr Tyr Lys Asp Cys (SEQ ID NO:12). Therefore, this exemplified Nef protein, Opt tPA-Nef, contains both a tPA leader sequence as well as deleting the myristylation site of Gly-2A DNA molecule encoding HIV-1 Nef from the HIV-1 jfrl isolate wherein the codons are optimized for expression in a mammalian system such as a human. 25

In another specific embodiment of the present invention, a DNA molecule is disclosed which encodes optimized HIV-1 Nef wherein the open reading frame of a recombinant adenoviral HIV vaccine encodes for modifications at the amino terminal myristylation site (Gly-2 to Ala-2) and substitution of the Leu-174-Leu-175 dileucine motif to Ala-174-Ala-175. This open reading frame is herein described as opt nef (G2A,LLAA) and is disclosed as SEQ ID NO:13, which comprises an initiating methionine residue at nucleotides 12-14 and a "TAA" stop codon from nucleotides 660-662. The nucleotide sequence of this codon optimized version of HIV-1 jrfl nef gene with the above mentioned modifications is disclosed herein as SEQ ID NO:13, as follows:

30

GATCTGCCAC CATGGCCGGC AAGTGGTCCA AGAGGTCCGT GCCCGGCTGG TCCACCGTGA GGGAGAGGAT GAGGAGGGCC GAGCCCGCCG GCGACAGGGT GAGGAGGACC GAGCCCGCCG GCGTGGGCGT GCCGACAGGCT GCGAGAAGCA CGGCGCCATC ACCTCCTCCA ACACCGCCGC CACCAACGCC GACTGCGCCT GGCTGGAGGC CCAGGAGGAC GAGGAGGTGG GCTTCCCCGT GAGGCCCCAG GTGCCCCTGA GGCCCATGAC CTACAAGGGC GCCGTGGACC TGTCCCACTT CCTGAAGGAG AAGGGCGGCC TGGAGGGCCT GATCCACTCC CAGAAGAGGC AGGACATCCT GGACCTGTGG GTGTACCACA CCCAGGGGCTA CTTCCCCGAC TGGCAGAACT ACACCCCCGG CCCCGGCATC AGGTTCCCCC TGACCTTCG CTGGTGCTTC AAGCTGGTGC CCGTGGAGCC CGAGAAGGGC GAGGAGGGCCA ACGAGGGCGA GAACAACTGC GCCGCCCACC CCATGTCCCC GAGAAGGGC AACAACTGC GCCGCCCACC CCATGTCCCC GCCGCCCACC GAGAGCGCC CCAGGGCGC CCGAGGAGCCC CCAAGCTGC CCCAGGCCC CCAAGCTGC CCCACCC CCAAGCTGC CTTCCACCAC CCCAAGCTGC CCCACCC CCAAGCCCC CGAGTACTAC AAGGACTGCT AAAGCCCGGG CC (SEQ ID NO:13).

The open reading frame of SEQ ID NO:13 encodes Nef (G2A,LLAA), disclosed herein as SEQ ID NO:14, as follows:

10

30

35

15 Met Ala Gly Lys Trp Ser Lys Arg Ser Val Pro Gly Trp Ser Thr Val Arg Glu Arg Met Arg Arg Ala Glu Pro Ala Ala Asp Arg Val Arg Arg Thr Glu Pro Ala Ala Val Gly Val Gly Ala Val Ser Arg Asp Leu Glu Lys His Gly Ala Ile Thr Ser Ser Asn Thr Ala Ala Thr Asn Ala Asp Cys Ala Trp Leu Glu Ala Gln Glu Asp Glu Glu Val Gly Phe Pro Val 20 Arg Pro Gln Val Pro Leu Arg Pro Met Thr Tyr Lys Gly Ala Val Asp Leu Ser His Phe Leu Lys Glu Lys Gly Gly Leu Glu Gly Leu Ile His Ser Gln Lys Arg Gln Asp Ile Leu Asp Leu Trp Val Tyr His Thr Gln Gly Tyr Phe Pro Asp Trp Gln Asn Tyr Thr Pro Gly Pro Gly Ile Arg Phe Pro Leu Thr Phe Gly Trp Cys Phe Lys Leu Val Pro Val Glu Pro 25 Glu Lys Val Glu Glu Ala Asn Glu Gly Glu Asn Asn Cys Ala Ala His Pro Met Ser Gln His Gly Ile Glu Asp Pro Glu Lys Glu Val Leu Glu Trp Arg Phe Asp Ser Lys Leu Ala Phe His His Val Ala Arg Glu Leu His Pro Glu Tyr Tyr Lys Asp Cys Ser (SEQ ID NO:14).

An additional embodiment of the present invention relates to another DNA molecule encoding optimized HIV-1 Nef wherein the amino terminal myristylation site and dileucine motif have been deleted, as well as comprising a tPA leader peptide. This DNA molecule, opt tpanef (LLAA) comprises an open reading frame which encodes a Nef protein containing a tPA leader sequence fused to amino acid residue 6-216 of HIV-1 Nef (jfrl), wherein Leu-174 and Leu-175 are substituted with Ala-174 and Ala-175 (Ala-195 and Ala-196 in this tPA-based fusion protein). The nucleotide

sequence comprising the open reading frame of opt tpanef (LLAA) is disclosed herein as SEQ ID NO:15, as shown below:

CATGGATGCA ATGAAGAGAG GGCTCTGCTG TGTGCTGCTG CTGTGTGGAG CAGTCTTCGT

TTCGCCCAGC GAGATCTCCT CCAAGAGGTC CGTGCCCGGC TGGTCCACCG TGAGGGAGAG

GATGAGGAGG GCCGAGCCCG CCGCCGACAG GGTGAGGAGG ACCGAGCCCG CCGCCGTGGG

CGTGGGCGCC GTGTCCAGGG ACCTGGAGAA GCACGGCGCC ATCACCTCCT CCAACACCGC

CGCCACCAAC GCCGACTGCG CCTGGCTGGA GGCCCAGGAG GACGAGGAGG TGGGCTTCCC

CGTGAGGCCC CAGGTGCCCC TGAGGCCCAT GACCTACAAG GGCGCCGTGG ACCTGTCCCA

CTTCCTGAAG GAGAAGGGCG GCCTGGAGGG CCTGATCCAC TCCCAGAAGA GGCAGGACAT

CCTGGACCTG TGGGTGTACC ACACCCAGGG CTACTTCCCC GACTGGCAGA ACTACACCCC

CGGCCCCGGC ATCAGGTTCC CCCTGACCTT CGGCTGGTGC TTCAAGCTGG TGCCCGTGGA

GCCCGAGAAG GTGGAGGAGG CCAACGAGGG CGAGAACAAC TGCGCCGCCC ACCCCATGTC

CCAGCACGGC ATCGAGGACC CCGAGAAGGA GGTGCTGGAG TGGAGGTTCG ACTCCAAGCT

GGCCTTCCAC CACGTGGCCA GGGAGCTGCA CCCCGAGTAC TACAAGGACT GCTAAAGCCC

(SEQ ID NO:15).

10

15

20

25

30

35

The open reading frame of SEQ ID NO:7 encoding tPA-Nef (LLAA), disclosed herein as SEQ ID NO:16, is as follows:

Met Asp Ala Met Lys Arg Gly Leu Cys Cys Val Leu Leu Leu Cys Gly Ala Val Phe Val Ser Pro Ser Glu Ile Ser Ser Lys Arg Ser Val Pro Gly Trp Ser Thr Val Arg Glu Arg Met Arg Arg Ala Glu Pro Ala Ala Asp Arg Val Arg Arg Thr Glu Pro Ala Ala Val Gly Val Gly Ala Val Ser Arg Asp Leu Glu Lys His Gly Ala Ile Thr Ser Ser Asn Thr Ala Ala Thr Asn Ala Asp Cys Ala Trp Leu Glu Ala Gln Glu Asp Glu Glu Val Gly Phe Pro Val Arg Pro Gln Val Pro Leu Arg Pro Met Thr Tyr Lys Gly Ala Val Asp Leu Ser His Phe Leu Lys Glu Lys Gly Gly Leu Glu Gly Leu Ile His Ser Gln Lys Arg Gln Asp Ile Leu Asp Leu Trp Val Tyr His Thr Gln Gly Tyr Phe Pro Asp Trp Gln Asn Tyr Thr Pro Gly Pro Gly Ile Arg Phe Pro Leu Thr Phe Gly Trp Cys Phe Lys Leu Val Pro Val Glu Pro Glu Lys Val Glu Glu Ala Asn Glu Gly Glu Asn Asn Cys Ala Ala His Pro Met Ser Gln His Gly Ile Glu Asp Pro Glu Lys Glu Val Leu Glu Trp Arg Phe Asp Ser Lys Leu Ala Phe His His Val Ala Arg Glu Leu His Pro Glu Tyr Tyr Lys Asp Cys (SEQ ID NO:16). An adenoviral vector of the present invention may comprise a DNA sequence, regardless of codon usage, which expresses a wild type or modified Nef protein as described herein, including but not limited to modified Nef proteins which comprise a deletion or substitution of Gly 2, a deletion of substitution of Leu 174 and Leu 175

and/or inclusion of a leader sequence. Therefore, partial or fully codon optimized DNA vaccine expression vector constructs are preferred since such constructs should result in increased host expression. However, it is within the scope of the present invention to utilize "non-codon optimized" versions of the constructs disclosed herein, especially modified versions of HIV Nef which are shown to promote a substantial cellular immune response subsequent to host administration.

Figure 20A-C show nucleotide sequences at junctions between nef coding sequence and plasmid backbone of nef expression vectors V1Jns/nef (Figure 20A), V1Jns/nef(G2A,LLAA) (Figure 20B), V1Jns/tpanef (Figure 20C) and V1Jns/tpanef(LLAA) (Figure 20C, also). 5' and 3' flanking sequences of codon optimized nef or codon optimized nef mutant genes are indicated by bold/italic letters; nef and nef mutant coding sequences are indicated by plain letters. Also indicated (as underlined) are the restriction endonuclease sites involved in construction of respective nef expression vectors. V1Jns/tpanef and V1Jns/tpanef(LLAA) have identical sequences at the junctions.

Figure 21 shows a schematic presentation of nef and nef derivatives. Amino acid residues involved in Nef derivatives are presented. Glycine 2 and Leucine 174 and 175 are the sites involved in myristylation and dileucine motif, respectively.

20

25

30

35

5

10

15

### **EXAMPLE 19**

#### MRKAd5Pol Construction and Virus Rescue

Construction of vector: shuttle plasmid and pre-adenovirus plasmid - Key steps performed in the construction of the vectors, including the pre-adenovirus plasmid denoted MRKAd5pol, is depicted in Figure 22. Briefly, the adenoviral shuttle vector for the full-length inactivated HIV-1 pol gene is as follows. The vector MRKpdelE1(Pac/pIX/pack450)+CMVmin+BGHpA(str.) is a derivative of the shuttle vector used in the construction of the MRKAd5gag adenoviral pre-plasmid. The vector contains an expression cassette with the hCMV promoter (no intronA) and the bovine growth hormone polyadenylation signal. The expression unit has been inserted into the shuttle vector such that insertion of the gene of choice at a unique BgIII site will ensure the direction of transcription of the transgene will be Ad5 E1 parallel when inserted into the MRKpAd5(E1-/E3+)Cla1 (or MRKpAdHVE3) preplasmid. The vector, similar to the original shuttle vector contains the Pac1 site, extension to the packaging signal region, and extension to the pIX gene. The synthetic full-length codon-optimized HIV-1 pol gene was isolated directly from the plasmid pV1Jns-HIV-pol-inact(opt). Digestion of this plasmid with BgI II releases the pol

gene intact (comprising a codon optimized IA pol sequence as disclosed in SEQ ID NO:3). The pol fragment was gel purified and ligated into the MRKpdelE1(Pac/pIX/pack450)+CMVmin+BGHpA(str.) shuttle vector at the BgIII site. The clones were checked for the correct orientation of the gene by using restriction enzymes DraIII/Not1. A positive clone was isolated and named MRKpdel+hCMVmin+FL-pol+bGHpA(s). The genetic structure of this plasmid was verified by PCR, restriction enzyme and DNA sequencing. The pre-adenovirus plasmid was constructed as follows. Shuttle plasmid MRKpdel+hCMVmin+FLpol+bGHpA(S) was digested with restriction enzymes Pac1 and Bst1107 I (or its isoschizomer, BstZ107 I) and then co-transformed into E. coli strain BJ5183 with linearized (Cla1 digested) adenoviral backbone plasmid, MRKpAd(E1-/E3+)Cla1. The resulting pre-plasmid originally named MRKpAd+hCMVmin+FLpol+bGHpA(S)E3+ is now referred to as "pMRKAd5pol". The genetic structure of the resulting pMRKAd5pol was verified by PCR, restriction enzyme and DNA sequence analysis. The vectors were transformed into competent E. coli XL-1 Blue for preparative production. The recovered plasmid was verified by restriction enzyme digestion and DNA sequence analysis, and by expression of the pol transgene in transient transfection cell culture. The complete nucleotide sequence of this pMRKAd5HIV-1pol adenoviral vector is shown in Figure 26 A-AO.

10

15

20

25

30

Generation of research-grade recombinant adenovirus - The pre-adenovirus plasmid, pMRKAd5pol, was rescued as infectious virions in PER.C6® adherent monolayer cell culture. To rescue infectious virus, 12  $\mu$ g of pMRKAd5pol was digested with restriction enzyme PacI (New England Biolabs) and 3.3  $\mu$ g was transfected per 6 cm dish of PER.C6® cells using the calcium phosphate coprecipitation technique (Cell Phect Transfection Kit, Amersham Pharmacia Biotech Inc.). PacI digestion releases the viral genome from plasmid sequences allowing viral replication to occur after entry into PER.C6® cells. Infected cells and media were harvested 6 -10 days post-transfection, after complete viral cytopathic effect (CPE) was observed. Infected cells and media were stored at  $\leq$  -60°C. This pol containing recombinant adenovirus is referred to herein as "MRKAd5pol". This recombinant adenovirus expresses an inactivated HIV-1 Pol protein as shown in SEQ ID NO:6.

#### **EXAMPLE 20**

### MRKAd5Nef Construction and Virus Rescue

35 Construction of vector: shuttle plasmid and pre-adenovirus plasmid - Key steps performed in the construction of the vectors, including the pre-adenovirus

plasmid denoted MRKAd5nef, is depicted in Figure 23. Briefly, as shown in Example 19 above, the vector

5

10

15

20

25

30

35

MRKpdelE1(Pac/pIX/pack450)+CMVmin+BGHpA(str.) is the shuttle vector used in the construction of the MRKAd5gag adenoviral pre-plasmid. It has been modified to contain the *Pac1* site, extension to the packaging signal region, and extension to the pIX gene. It contains an expression cassette with the hCMV promoter (no intronA) and the bovine growth hormone polyadenylation signal. The expression unit has been inserted into the shuttle vector such that insertion of the gene of choice at a unique *Bgl11* site will ensure the direction of transcription of the transgene will be Ad5 E1 parallel when inserted into the MRKpAd5(E1-/E3+)Cla1 pre-plasmid. The synthetic full-length codon-optimized HIV-1 nef gene was isolated directly from the plasmid pV1Jns/nef (G2A,LLAA). Digestion of this plasmid with *Bgl11* releases the pol gene intact, which comprises the nucleotide sequence as disclosed in SEQ ID NO:13. The nef fragment was gel purified and ligated into the

MRKpdelE1+CMVmin+BGHpA(str.) shuttle vector at the Bgl11 site. The clones were checked for correction orientation of the gene by using restriction enzyme Scal. A positive clone was isolated and named MRKpdelE1hCMVminFL-nefBGHpA(s). The genetic structure of this plasmid was verified by PCR, restriction enzyme and DNA sequencing. The pre-adenovirus plasmid was constructed as follows. Shuttle plasmid MRKpdelE1hCMVminFL-nefBGHpA(s) was digested with restriction enzymes Pac1 and Bst1107 I (or its isoschizomer, BstZ107 I) and then co-transformed into E. coli strain BJ5183 with linearized (Cla1 digested) adenoviral backbone plasmid, MRKpAd(E1/E3+)Cla1. The resulting pre-plasmid originally named MRKpdelE1hCMVminFL-nefBGHpA(s) is now referred to as "pMRKAd5nef". The genetic structure of the resulting pMRKAd5nef was verified by PCR, restriction enzyme and DNA sequence analysis. The vectors were transformed into competent E. coli XL-1 Blue for preparative production. The recovered plasmid was verified by restriction enzyme digestion and DNA sequence analysis, and by expression of the nef transgene in transient transfection cell culture. The complete nucleotide sequence of this pMRKAd5HIV-1nef adenoviral vector is shown in Figure 27A-AM.

Generation of research-grade recombinant adenovirus - The pre-adenovirus plasmid, pMRKAd5nef, was rescued as infectious virions in PER.C6<sup>®</sup> adherent monolayer cell culture. To rescue infectious virus, 12 μg of pMRKAdnef was digested with restriction enzyme Pac1 (New England Biolabs) and 3.3 μg was transfected per 6 cm dish of PER.C6<sup>®</sup> cells using the calcium phosphate coprecipitation technique (Cell Phect Transfection Kit, Amersham Pharmacia Biotech

Inc.). Pac1 digestion releases the viral genome from plasmid sequences allowing viral replication to occur after entry into PER.C6®cells. Infected cells and media were harvested 6-10 days post-transfection, after complete viral cytopathic effect (CPE) was observed. Infected cells and media were stored at  $\leq$  -60°C. This nef containing recombinant adenovirus is now referred to as "MRKAd5nef".

### EXAMPLE 21

5

10

15

20

25

30

35

# Construction of Murine CMV Promoter Containing Shuttle Vectors for Inactivated Pol and Nef/G2A,LLAA

The murine CMV (mCMV) was amplified from the plasmid pMH4 (supplied by Frank Graham, McMaster University) using the primer set: mCMV (Not I) Forward: 5'-ATA AGA ATG CGG CCG CCA TAT ACT GAG TCA TTA GG-3' (SEQ ID NO: 20); mCMV (Bgl II)Reverse: 5'-AAG GAA GAT CTA CCG ACG CTG GTC GCG CCT C-3' (SEQ ID NO:21). The underlined nucleotides represent the Not I and the Bgl II sites respectively for each primer. This PCR amplicon was used for the construction of the mCMV shuttle vector containing the transgene in the E1 parallel orientation. The hCMV promoter was removed from the original shuttle vector (containing the hCMV-gag-bGHpA transgene in the E1 parallel orientation) by digestion with Not I and Bgl II. The mCMV promoter (Not I/Bgl II digested PCR product) was inserted into the shuttle vector in a directional manner. The shuttle vector was then digested with Bgl II and the gag reporter gene (Bgl II fragment) was re-inserted back into the shuttle vector. Several clones were screened for correct orientation of the reporter gene. For the construction of the mCMV-gag in the E1 antiparallel orientation, the mCMV promoter was amplified from the plasmid pMH4 using the following primer set: mCMV (Asc I) Forward: 5'- ATA AGA ATG GCG CGC CAT ATA CTG AGT CAT TAG G (SEQ ID NO:22); mCMV (Bgl II) Reverse: 5' AAG GAA GAT CTA CCG ACG CTG GTC GCG CCT C (SEQ ID NO:23). The underlined nucleotides represent the Asc I and Bgl II sites, respectively for each primer. The shuttle vector containing the hCMV-gag transgene in the E1 antiparallel orientation was digested with Asc1 and Bgl11 to remove the hCMV-gag portion of the transgene. The mCMV promoter (Asc1/Bgl11 digested PCR product) was inserted into the shuttle vector in a directional manner. The vector was then digested with Bgl11 and the gag reporter gene (Bgl11 fragment) was re-inserted. Several clones were screened for correct orientation of the reporter gene. For each of the full length IA pol and full length nef/G2A,LLAA genes, cloning was performed using the unique

 $Bgl ext{ II}$  site within the mCMV-bGHpA shuttle vector. The pol and nef genes were excised from their respective pV1Jns plasmids by  $Bgl ext{ II}$  digestion.

#### **EXAMPLE 22**

Construction of mCMV Full Length Inactivated Pol and Full Length nef/G2A.LLAA Adenovectors

5

10

15

30

Each of these transgenes of Example 21 were inserted into the modified shuttle vector in both the E1 parallel and E1 anti-parallel orientations. *Pac1* and *BstZ110I* digestion of each shuttle vector was performed and each specific transgene fragment containing the flanking Ad5 sequences was isolated and co-transformed with *Cla I* digested MRKpAd5(E3+) or MRKpAd5(E3-) adenovector plasmids via bacterial homologous recombination in BJ5183 *E. coli* cells. Recombinant preplasmid adenovectors containing the various transgenes in both the E3- and E3+ versions (and in the E1 parallel and E1 antiparallel orientations) were subsequently prepared in large scale following transformation into XL-1 Blue *E. coli* cells and analyzed by restriction analysis and sequencing.

#### **EXAMPLE 23**

Construction of hCMV-tpa-nef (LLAA) Adenovector

The tpa-nef gene was amplified out from GMP grade pV1Jns-tpanef (LLAA) vector using the primer sets: Tpanef (BamHI) F 5'-ATT GGA TCC ATG GAT GCA ATG AAG AGA GGG (SEQ ID 24); Tpanef (BamHI) R 5'-ATA GGA TCC TTA GCA GTC CTT GTA GTA CTC G (SEQ ID NO:25). The resulting PCR product was digested with BamHI, gel purified and cloned into the Bgl II site of MRKAd5CMV-bGHpA shuttle vector (Bgl II digested and calf intestinal phosphatase treated). Clones containing the tpanef (LLAA) gene (see SEQ ID NO:15 for complet coding region) in the correct orientation with respect to the hCMV promoter were selected following Sca I digestion. The resulting MRKAd5tpanef shuttle vector was digested with Pac I and Bst Z1101 and cloned into the E3+ MRKAd5 adenovector via bacterial

#### **EXAMPLE 24**

Immunogenicity of MRKAd5pol and MRKAd5nef Vaccine

Materials and Methods - Rodent Immunization - Groups of N=10 BALB/c

mice were immunized i.m. with the following vectors: (1) MRKAd5hCMV-IApol

(E3+) at either 10^7 vp and 10^9 vp; and (2) MRKAd5hCMV-IApol (E3-) at either

homologous recombination techniques.

10^7 vp and 10^9 vp. At 7 weeks post dose, 5 of the 10 mice per cohort were boosted with the same vector and dose they initially received. At 3 weeks post the second does, sera and spleens were collected from all the animals for RT ELISA and IFNg ELIspot analyses, respectively. For all rodent immunizations, the Ad5 vectors were diluted in 5 mM Tris, 5% sucrose, 75 mM NaCl, 1 mM MgCl2, 0.005% polysorbate 80, pH 8.0. The total dose was injected to both quadricep muscles in 50 μL aliquots using a 0.3-mL insulin syringe with 28-1/2G needles (Becton-Dickinson, Franklin Lakes, NJ).

Groups of N=10 C57/BL6 mice were immunized i.m. with the following vectors: (1) MRKAd5hCMV-nef(G2A,LLAA) (E3+) at either 10^7 vp and 10^9 vp; (2) MRKAd5mCMV-nef(G2A,LLAA) (E3+) at either 10^7 vp and 10^9 vp; and (3) MRKAd5mCMV-tpanef(LLAA) (E3+) at either 10^7 vp and 10^9 vp. At 7 weeks post dose, 5 of the 10 mice per cohort were boosted with the same vector and dose they initially received. At 3 weeks post the second does, sera and spleens were collected from all the animals for RT ELISA and IFNg ELIspot analyses, respectively.

Non-human Primate immunization - Cohorts of 3 rhesus macaques (2-3 kg) were vaccinated with the following Ad vectors: (1) MRKAd5hCMV-IApol (E3+) at either 10^9 vp and 10^11 vp dose; and (2) MRKAd5hCMV-IApol (E3-) at either 10^9 vp and 10^11 vp; (3) MRKAd5hCMV-nef(G2A,LLAA) (E3+) at either 10^9 vp and 10^11 vp; and (4) MRKAd5mCMV-nef(G2A,LLAA) (E3+) at either 10^9 vp and 10^11 vp. The vaccine was administered to chemically restrained monkeys (10 mg/kg ketamine) by needle injection of two 0.5 mL aliquots of the Ad vectors (in 5 mM Tris, 5% sucrose, 75 mM NaCl, 1 mM MgCl<sub>2</sub>, 0.005% polysorbate 80, pH 8.0) into both deltoid muscles. The animals were immunized twice at a 4 week interval (T=0, 4 weeks).

Murine anti-RT and anti-nef ELISA - Anti-RT titers were obtained following standard secondary antibody-based ELISA. Maxisorp plates (NUNC, Rochester, NY) were coated by overnight incubation with 100 μL of 1 μg/mL HIV-1 RT protein (Advanced Biotechnologies, Columbia, MD) in PBS. For anti-nef ELISA, 100 uL of 1 ug/mL HIV-1 nef (Advanced Biotechnologies, Columbia, MD) was used to coat the plates. The plates were washed with PBS/0.05% Tween 20 using Titertek MAP instrument (Hunstville, AL) and incubated for 2 h with 200 μL/well of blocking solution (PBS/0.05% tween/1% BSA). An initial serum dilution of 100-fold was performed followed by 4-fold serial dilution. 100-μL aliquots of serially diluted samples were added per well and incubated for 2 h at room temperature. The plates

were washed and 100  $\mu$ L of 1/1000-diluted HRP-rabbit anti-mouse IgG (ZYMED, San Francisco, CA) were added with 1 h incubation. The plates were washed thoroughly and soaked with 100  $\mu$ L 1,2-phenylenediamine dihydrochloride/hydrogen peroxide (DAKO, Norway) solution for 15 min. The reaction was quenched by adding 100  $\mu$ L of 0.5M H<sub>2</sub>SO4 per well. OD<sub>492</sub> readings were recorded using Titertek Multiskan MCC/340 with S20 stacker. Endpoint titers were defined as the highest serum dilution that resulted in an absorbance value of greater than or equal to 0.1 OD<sub>492</sub> (2.5 times the background value).

5

25

30

35

Non-human primate and murine ELIspot assays - The enzyme-linked immuno-spot (ELISpot) assay was utilized to enumerate antigen-specific INFy-10 secreting cells from mouse spleens (Miyahira, et al. 1995, J. Immunol. Methods 181:45-54) or macaque PBMCs. Mouse spleens were pooled from 5 mice/cohort and single cell suspensions were prepared at 5x10<sup>6</sup>/mL in complete RPMI media (RPMI1640, 10% FBS, 2mM L-glutamine, 100U/mL Penicillin, 100 u/mL 15 streptomycin, 10 mM Hepes, 50 uM β-ME). Rhesus PBMCs were prepared from 8-15 mL of heparinized blood following standard Ficoll gradient separation (Coligan, et al, 1998, Current Protocols in Immunology. John Wiley & Sons, Inc.). Multiscreen opaque plates (Millipore, France) were coated with 100 µL/well of either 5 µg/mL purified rat anti-mouse IFN-y IgG1, clone R4-6A2 (Pharmingen, San Diego, CA), or 20 15 ug/mL mouse anti-human IFN-γ IgG<sub>2a</sub> (Cat. No. 1598-00, R&D Systems, Minneapolis, MN) in PBS at 4°C overnight for murine or monkey assays. respectively. The plates were washed with PBS/penicillin/streptomycin and blocked with 200 µL/well of complete RPMI media for 37 °C for at least 2 h.

To each well, 50 μL of cell samples (4-5x10<sup>5</sup> cells per well) and 50 μL of the antigen solution were added. To the control well, 50 μL of the media containing DMSO were added; for specific responses, either selected peptides or peptide pools (4 ug/mL per peptide final concentration) were added. For BALB/c mice immunized with the pol constructs, stimulation was conducted using a pool of CD4<sup>+</sup>-epitope containing 20-mer peptides (aa21-40, aa411-430, aa641-660, aa731-750, aa771-790) or a pool of CD8<sup>+</sup>-epitope containing peptides (aa201-220, aa311-330, aa781-800). For C57/BL6 mice immunized with the nef construct, either aa51-70 (CD8<sup>+</sup> T cell epitope) or aa81-100 (CD4<sup>+</sup>) peptide derived from the nef sequence was added for specific stimulation. In monkeys, the responses against pol were evaluated using two pools (L and R) of 20-aa peptides that encompass the entire pol sequence and overlap by 10 amino acids. In monkeys vaccinated with the nef constructs, a single pool containing 20-mer peptides covering the entire HIV-1 nef sequence and overlapping

by 10 aa was used. Each sample/antigen mixture was performed in triplicate wells for murine samples or in duplicate wells for rhesus PBMCs. Plates were incubated at 37°C, 5% CO<sub>2</sub>, 90% humidity for 20-24 h. The plates were washed with PBS/0.05% Tween 20 and incubated with 100 μL/well of either 1.25 μg/mL biotin-conjugated rat anti-mouse IFN-γ mAb, clone XMG1.2 (Pharmingen) or of 0.1 ug/mL biotinylated anti-human IFN-gamma goat polyclonal antibody (R&D Systems) at 4°C overnight. The plates were washed and incubated with 100 μL/well 1/2500 dilution of strepavidin-alkaline phosphatase conjugate (Pharmingen) in PBS/0.005% Tween/5% FBS for 30 min at 37 °C. Spots were developed by incubating with 100 μL/well 1-step NBT/BCIP (Pierce Chemicals) for 6-10 min. The plates were washed with water and allowed to air dry. The number of spots in each well was determined using a dissecting microscope and the data normalized to 10<sup>6</sup> cell input.

Non-human Primate anti-RT ELISA - The pol-specific antibodies in the monkeys were measured in a competitive RT EIA assay, wherein sample activity is determined by the ability to block RT antigen from binding to coating antibody on the plate well. Briefly, Maxisorp plates were coated with saturating amounts of pol positive human serum (#97111234). 250 uL of each sample is incubated with 15 uL of 266 ng/mL RT recombinant protein (in RCM 563, 1% BSA, 0.1% tween, 0.1% NaN<sub>3</sub>) and 20 uL of lysis buffer (Coulter p24 antigen assay kit) for 15 min at room temperature. Similar mixtures are prepared using serially diluted samples of a standard and a negative control which defines maximum RT binding. 200 uL/well of each sample and standard were added to the washed plate and the plate incubated 16-24 h at room temperature. Bound RT is quantified following the procedures described in Coulter p24 assay kit and reported in milliMerck units per mL arbitrarily defined by the chosen standard.

Results - Rodent Studies - BALB/c mice (n=5 mice/cohort) were immunized once or twice with varying doses of MRKAd5hCMV-IApol(E3+) and MRKAd5hCMV-IApol(E3-). At 3 weeks after the second dose, Anti-pol IgG levels were determined by an ELISA assay using RT as a surrogate antigen. Cellular response were quantified via IFNy ELISpot assay against pools of pol-epitope containing peptides. The results of these assays are summarized in Table 10. The results indicate that the mouse vaccinees exhibited detectable anti-RT IgGs with an adenovector dose as low as 10^7 vp. The humoral responses are highly dose-dependent and are boostable with a second immunization. One or two doses of either pol vectors elicit high frequencies of antigen-specific CD4<sup>+</sup> and CD8<sup>+</sup> T cells; the responses are weakly dose-dependent but are boostable with a second immunization.

Table 10. Immunogenicity of MRKAd5pol Vectors in BALB/c mice.

|       |                       |          |                 | Ап   | ti-RT IgG Tite  | rs*            | S             | FC/10^6 cel             | ls°                  |
|-------|-----------------------|----------|-----------------|--|-----------------|----------------|---------------|-------------------------|----------------------|
| Group | Vaccine               | I Dose I | No. of<br>Doses | GMT  | +SE             | -SE            | Medium        | CD4+<br>peptide<br>pool | CD8+<br>peptide      |
| 1     | MRKAd5hCMVFLpol (E3+) | 10^7 vp  | 2<br>1          | 310419<br>919                                | 301785<br>372   | 153020<br>265  | 1(1)<br>1(1)  | 75(4)<br>72(9)          | 2313(67)<br>533(41)  |
| 2     | MRKAd5hCMVFLpol (E3+) | 10^9 vp  | 2               | 1638400 <sup>b</sup><br>713155               | 0<br>528520     | 0<br>303555    | 2(2)<br>1(1)  | 114(9)<br>48(7)         | 2063(182)<br>733(89) |
| 3     | MRKAd5hCMVFLpol (E3-) | 10^7 vp  | 2               | 310419<br>6400                               | 386218<br>14013 | 172097<br>4393 | 0(0)<br>10(8) | 223(7)<br>141(21)       | 2607(27)<br>409(28)  |
| 4     | MRKAd5hCMVFLpol (E3-) | 10^9 vp  | 2               | 1638400 <sup>b</sup><br>1241675 <sup>b</sup> | 0<br>396725     | 0<br>300661    | 1(1)<br>0(0)  | 160(13)<br>39(13)       | 2385(11)<br>833(83)  |
| 5     | Naïve                 | none     | none            | 57   | 9               | 7              | 9(2)          | 11(4)                   | 10(1)                |

\*GMT, geometric mean titler of the cohort of 5 mice; SE, standard error of the gemetric mean

5 C57/BL6 mice were immunized once or twice with varying doses of MRKAd5hCMV-nef(G2A,LLAA) (E3+), MRKAd5mCMV-nef(G2A,LLAA) (E3+) at either 10^7 vp and(3) MRKAd5mCMV-tpanef(LLAA) (E3+) at either 10^7 vp and 10^9 vp. The immune response were analyzed using similar protocols and the results are listed in Table 11. While anti-nef IgG responses could not be detected in this model system with any of the constructs, there are strong indications of a cellular immunity generated against nef using the ELIspot assay.

Table 11. Immunogenicity of MRKAd5nef Vectors in C57/BL6 mice.

|       |                       |         |                 | An  | ti-nef IgG Tite | rs" | S       | FC/10^6 cell    | s <sup>b</sup>   |
|-------|-----------------------|---------|-----------------|-----|-----------------|-----|---------|-----------------|------------------|
| Group | Vaccine               | Dose    | No. of<br>Doses | GMT | +SE             | -SE | Medium  | aa51-70<br>CD8+ | aa81-100<br>CD4+ |
| 1     | MRKAd5hCMVFLnef (E3+) | 10^7 vp | 2               | 174 | 70              | 50  | 1(1)    | 23(1)           | 1(1)             |
| , i   |                       |         | 1               | 132 | 42              | 32  | 0(0)    | 0(0)            | 0(0)             |
| 2.    | MRKAd5hCMVFLnef (E3+) | 10^9 vp | 2               | 174 | 70              | 50  | 0(0)    | 61(7)           | 4(2)             |
|       | •                     |         | 1               | 132 | 42              | 32  | 1(1)    | 62(7)           | 3(1)             |
| 3     | MRKAd5mCMVFLnef (E3+) | 10^7 vp | 2               | 132 | 42              | 32  | 3(1)    | 15(5)           | 5(2)             |
|       |                       |         | 1               | 115 | 46              | 33  | 3(2)    | 3(2)            | 4(2)             |
| 4     | MRKAd5mCMVFLnef (E3+) | 10'9 vp | 2               | 132 | 42              | 32  | 4(2)    | 83(13)          | 5(1)             |
|       |                       |         | 1               | 132 | 42              | 32  | 2(1)    | 29(2)           | 4(0)             |
| 5     | MRKAd5mCMVtpanef(E3+) | 10^7 vp | 2               | 132 | 42              | 32  | 3(2)    | 14(2)           | 5(1)             |
|       | ·                     |         | 1               | 100 | 0               | 0   | 3(1)    | 13(4)           | 10(3)            |
| 6     | MRKAd5mCMVtpanef(E3+) | 10^9 vp | 2               | 230 | 170             | 98  | 3(2)    | 145(29)         | 4(0)             |
|       |                       | 1       | 1               | 115 | 46              | 33  | 7(1)    | 151(14)         | 10(0)            |
| 7     | Naīve                 | none    | none            | 152 | 78              | 52  | - 21(2) | · 18(6)         | 26(3)            |

\*GMT, geometric mean titer of the cohort of 5 mice; SE, standard error of the gemetric mean

15

Monkey Studies - Cohorts of 3 rhesus macaques were immunized with 2 doses of MRKAd5hCMV-IApol(E3+) and MRKAd5hCMV-IApol(E3-). The number of antigen-specific T cells (per million PBMCs) were enumerated using one of two

<sup>&</sup>lt;sup>b</sup>Near or at the upper limit of the serial dilution; hence, could be greater than this value

No. of Spot-forming Cells per million spleonoytes; mean values of triplicates are reported along with standard errors in parenthesis.

No. of spot-forming cells per million spleonoytes; mean values of triplicates are reported along with standard errors in parenthesis.

peptide pools (L and R) that cover the entire pol sequence; the results are listed in Table 12. Moderate-to-strong T cell responses were detected in the vaccinees using either constructs even at a low dose of 10^9 vp. Longitudinal analyses of the anti-RT antibody titers in the animals suggest that the pol transgene product is expressed efficiently to elicit a humoral response (Table 13). It would appear that generally higher immune responses were observed in animals that received the E3- construct compared to the E3+ virus.

Table 12. Pol-specific T Cell Responses in MRKAd5pol Immunized Rhesus

10 Macaques.

5

| Vaccine (T=0,4 wks)    | Monk # |      | Prebleec | 1     |      | T=4   |       |      | T=7   |       |      | T=16  |       |
|------------------------|--------|------|----------|-------|------|-------|-------|------|-------|-------|------|-------|-------|
|                        |        | Mock | Pol L    | Pol R | Mock | Pol L | Pol R | Mock | Pal L | Pol R | Mock | Pol L | Pol R |
| MRK Ad5hCMV-lApot(E3+) | 99C100 | 1    | 0        | 0     | 1    | 38    | 31    | 0    | 52    | 146   | ٥    | 49    | 715   |
| 10411 vp               | 99C215 | 1 1  | 2        | 2     | 10   | 98    | 249   | 1    | 109   | 305   | 22   | 88    | 250   |
| 10 11 4p               | 99D201 | 5    | 5        | 4     | 6    | 149   | 95    | 0    | 40    | 35    | 0    | 35    | 18    |
| MRKAd5hCMV-IApol(E3+)  | 99D212 | 0    | 2        | 0     | 4    | 331   | 114   | 0    | 58    | 14    | 0    | 6     | 6     |
| 10/9 vp                | 99D180 | 0    | 4        | 2     | 0    | 19    | 192   | 4    | 36    | 156   | 5    | 38    | 106   |
|                        | 990201 | 8    | 5        | 21    | 6    | 62    | 62    | 0    | 18    | 32    | ۱۱   | 14    | 65    |
| MRKAd5hCMV-IApal(E3-)  | 99D239 | 5    | 2        | 2     | 20   | 82    | 172   | 1    | 66    | 114   | 9    | 21    | 40    |
| 10/11 vp               | 99C186 | 4    | 12       | 6     | 5    | 120   | 421   | 2    | 271   | 489   | 16   | 875   | 530   |
|                        | 99C084 | 1    | 8        | 9     | В    | 84    | 464   | 0    | 14    | 236   | 1    | 24    | 264   |
| MRKAd5hCMV-IApol(E3-)  | CC7C   | 10   | 10       | 8     | 12   | 724   | 745   | 4    | 322   | 376   | 4    | 188   | 176   |
| 10/9 vp                | CDIG   | 2    | 1 0      | 1 1   | 5    | 474   | 468   | 0    | 232   | 212   | 0    | 101   | 121   |
|                        | CDII   | 6    | 6        | 12    | 10   | 98    | 110   | 5    | 60    | 80    | 8    | 25    | 34    |
| Naive                  | 0830   | nd   | nd       | nd    | nd   | nd    | nd    | 4    | 2     | 2     | 2    | 1     | 2     |

nd, not determined Reported are SFC per million PBMCs; mean of duplicate wells.

Table 13. Anti-RT Ig Levels in MRKAd5pol Immunized macaques.

| RT ANTIBODY ASSAYTITERS IN mMU/ | mL  |         |      |      |
|---------------------------------|-----|---------|------|------|
| Vaccine/Monkey Tag              | T=4 | T=7     | T=12 | T=16 |
| MRKAd5hCMV-IApol(E3+), 10^11 vp |     |         |      |      |
| 99C100                          | 61  | 1999    | 5928 | 4768 |
| 99C215                          | 81  | 1541    | 2356 | 2767 |
| 99D201                          | 53  | 336     | 539  | 387  |
| MRKAd5hCMV-IApol(E3+), 10^9 vp  |     |         |      |      |
| 99D212                          | 10  | 40      | 49   | 68   |
| 99D180                          | <10 | 36      | 79   | 93   |
| 99C201                          | <10 | 37      | 71   | 76   |
| MRKAd5hCMV-IApol(E3-), 10^11 vp |     |         |      |      |
| 99D239                          | 44  | 460     | 1234 | 1015 |
| 99C186                          | 21  | · 233 · | 480  | 345_ |
| 990084                          | 235 | 2637    | 2858 | 1626 |
| MRKAd5hCMV-IApol(E3-), 10^9 vp  |     |         |      |      |
| CC7C                            | 32  | 175     | 306  | 235  |
| മാഭ                             | 20  | 140     | 273  | 419  |
| <b>©</b>                        | 15  | 112     | 149  | 237  |
|                                 | _   |         |      |      |

When rhesus macaques were immunized i.m. with two doses of MRKAd5nef

5 constructs, vigorous T cell responses ranging from 100 to as high as 1100 per million were observed in 8 of 12 vaccinees (Table 14). The efficacies of the mCMV- and hCMV- driven nef constructs are comparable on the basis of the data generated thus far.

10 Table 14. Nef-specific T cell Responses in MRKAd5nef Immunized Rhesus Macagues.

| Vaccine (T=0,4 wks)            | Monk #   | P    | re  | T    | =4    | T:   | <del>-</del> 7 | T=     | :16  |
|--------------------------------|----------|------|-----|------|-------|------|----------------|--------|------|
|                                |          | Mock | Nef | Mock | Nef   | Mock | Nef            | Mock   | Nef  |
| MRKAd5hCMV-nef(G2A,LLAA) (E3+) | CD2D     | 0    | 4   | 31   | 440   | 4    | 368            | 1      | 251  |
| 10^11 vp                       | CC7B     | 0    | 0   | 2    | 521   | 0    | 178            | 1      | 1522 |
|                                | CC61     | 2    | 9   | 31   | 112   | 0    | 108            | 11     | 100  |
| MRKAd5hCMV-nef(G2A,LLAA) (E3+) | CC2K     | 9    | 9   | 6    | 52    | 0    | 35             | 0      | 15   |
| 10^9 vp                        | CD15     | 5    | 4   | 30   | 998   | 2    | 586            | 0      | 434  |
|                                | CD16     | -6   | 1   | 6    | 1146  | 0    | 369            | 1 1    | 212  |
| MRKAd5mCMV-nef(G2A,LLAA) (E3+) | 99D191   | 1    | 5   | 4    | 614   | 0    | 298            | 2      | 419  |
| 10^11 vp                       | 99D144   | 4    | 6   | 5    | 434   | 0 '  | 1100           | 2      | 932  |
|                                | 99C193   | 1    | 2   | 1 1  | 58    | 1    | 22             | 0      | 64   |
| MRKAd5mCMV-nef(G2A,LLAA) (E3+) | - 99D224 | 1.   | 11  | 14   | _ 231 | . 1  | 125            | ·- · 0 | 70   |
| 10^9 vp                        | 99D250   | 8    | 9   | 4    | 108   | 0    | 54             | 0      | 5    |
|                                | 99C120   | 1    | 6   | 20   | 299   | 0    | 92             | 0      | 79   |
| Naîve                          | 083Q     | nd   | nd  | 18   | 22    | 4    | 5              | 2      | 1    |

### **EXAMPLE 25**

15 Comparison of Clade B vs. Clade C T Cell Responses in HIV-Infected Subjects PBMC samples collected from two dozens of patients infected with HIV-1 in US were tested in ELISPOT assays with peptide pools of 20-mer peptides overlapping by 10 amino acids. Four different peptide pools were tested for cross-clade recognition, and they were either derived from a clade B-based isolate (gag H-b; nefb) or a clade C-based isolate (gag H-c, nef-c). Data in Table 15 shows that T cells 20 from these patients presumably infected with clade B HIV-1 could recognize clade C gag and nef antigens in ELISPOT assay. Correlation analysis further demonstrated that these T cell responses against clade C gag peptide pool were about 60% of the clade B counterpart (Figure 24), while the T cell responses against clade C nef were about 85% of the clade B counterpart (Figure 25). These results suggest that cellular 25 immune responses generated in patients infected with clade B HIV-1 can recognize gag and nef antigens derived from clade C HIV-1. These data show that a HIV vaccine, such as a DNA or MRKAd5-based adenoviral vaccine expressing a clade B

gag and/or nef antigen will potentially have the ability to provide a prophylactic and/or therapetic advantage on a global scale.

5

15

20

Table 15
Responses Shown as the Number of gIFN-Secreting T Cells per Million PBMCs

| subject | bleed date | gag epitope # | mock | gag H-b | gagH-c | nef-b | nef-c |
|---------|------------|---------------|------|---------|--------|-------|-------|
|         | (          | from mapping) |      |         |        |       |       |
| #100    | 19-Jul-99  | 12            | 10   | 3950    | 1385   | 1295  | 1300  |
| #101    | 25-Jul-99  | 3             | 15   | 3885    | 1280   | na    | 1020  |
| #102    | 25-Jul-99  | 4             | 15   | 1740    | 850    | 1255  | 1785  |
| #104    | 7-Jun-99   | 2             | 5    | 1355    | 1185   | na    | 1060  |
| #107    | 11-Oct-99  | 2             | 25   | 3305    | 2795   | 670   | 870   |
| #405    | 11-Jul-99  | 2             | 15   | 4575    | 3180   | 1700  | 1500  |
| #501    | 19-Jul-99  | 2             | 15   | 1100    | 570    | 3365  | 3460  |
| #505    | 18-Jul-99  | 5             | 10   | 2145    | 1725   | 1235  | na    |
| #506    | 28-Feb-99  | 2             | 25   | 150     | 45     | 400   | 610   |
| #701    | 28-Mar-99  | 5             | 30   | 7620    | 4775   | 3320  | 2780  |
| #709    | 17-May-99  | 3             | 15   | 2785    | 1945   | 1090  | 1630  |
| #710    | 24-May-99  |               | 5    | 1055    | 1080   | 2210  | 2140  |
|         |            |               |      |         |        |       |       |

10 EXAMPLE 26

Characterization and Production of MRKAd5pol and MRKAd5nef

Vectors in Roller Bottles

Expansion of nef and pol Adenovectors - Nef and pol CsCl purified MRKAd5 seeds were used to infect roller bottles to produce P4 virus to be used as a seed for further experiments. P4 MRKAd5 pol and nef vectors were used to infect roller bottles at an MOI 280 vp/cell, except for hCMV-tpa-nef [E3+] which was infected at an MOI of 125 due to low titers of seed obtained at P4.

Table 16 Viral particle concentrations for P5 nef and pol adenovectors

| Adenovector        | AEX Titer (10 <sup>10</sup> vp/ml culture) | AEX Titer (10 <sup>4</sup> vp/cell) | Amplification<br>Ratio |
|--------------------|--|-------------------------------------|------------------------|
| hCMV-FL-nef [E3+]  | 1.1  | 0.9                                 | . 30                   |
| mCMV-FL-nef [E3+]  | 2.2  | 2.1                                 | 75                     |
| hCMV-tpa-nef [E3+] | 0.07                                       | 0.1                                 | 5                      |
| mCMV-tpa-nef [E3+] | 1.3  | 0.9                                 | 35                     |
| hCMV-FL-pol [E3+]  | 2.7  | 2.1                                 | 75                     |
| hCMV-FL-pol [E3-]  | 1.9  | 1.3                                 | 45                     |

5 Roller Bottle Passaging - Passaging of the pol and nef constructs continued through passage seven. Cell-associated (freeze/thaw lysis) and whole broth (tritonlysis) titers obtained in all passages were very consistent. In general, MRKAd5pol is ca. 70% as productive as MRKAd5gag while MRKAd5nef is ca. 25% as productive as MRKAd5gag. Samples of P7 virus for both constructs were analyzed by V&CB by restriction digest analysis and did not show any rearrangements.

Table 17. Passage Six Viral Productivity for MRKAd5pol and MRKAd5pef

|                   | 1    | Xviable (10° cells/ml),<br>Viability (%) |           | Cell Passage | AEX Titer<br>(Cell Associated) | Titer       | Amplification | Triton Lysis Titer |
|-------------------|------|--|-----------|--------------|--------------------------------|-------------|---------------|--------------------|
|                   | 1    | Infection                                | Harvest   | Number       | 1010 vp/mi culture             | 10° vp/cell | Ratio         | 1010 vp/ml culture |
| hCMV-PL-nef [B3+] | pool | 1.22, 85%                                |           | 62           | 0.8                            | 0.7         | 25            | 1.6                |
|                   | 1    |  | 0.99, 62% |              |                                |             |               |                    |
|                   | 2    |  | 1.10, 72% | }            | \ \ \                          |             | 1             |                    |
| bCMV-FL-pol [E3+] | pool | 1.42, 89%                                |           | 62           | 4.5                            | 3.2         | 115           | 7.0                |
|                   | 1    |  | 1.22, 70% |              |                                |             |               |                    |
|                   | 2    |  | 1.42, 74% |              |                                |             |               |                    |

15 Table 18. Passage Seven Viral Productivity for MRKAd5pol and MRKAd5nef

|                     |      |            | 0 <sup>6</sup> cells/ml),<br>ity (%)<br>Harvest | Cell Passage<br>Number | AEX Titer<br>(Cell Associated)<br>10 <sup>70</sup> vp/ml culture | Titer  10 <sup>4</sup> vp/ceii | Amplification<br>Ratio | Triton Lysis Titer             |
|---------------------|------|------------|---|------------------------|--|--------------------------------|------------------------|--------------------------------|
| hCMV-FL-nef [E3+]   |      |            | Tiarvest  |                        |  |                                |                        | 10 <sup>10</sup> vp/ml culture |
| ICW A-LT-USI [173+] | Pool | 1.33, 90%  | _   | 66                     | 1.0  | 0.8                            | 29                     | 2.1                            |
|                     | 1    |            | 0.96, 70%                                       | 1                      |  |                                |                        |                                |
|                     | 2    |            | 1.18, 73%                                       | .                      | 1  |                                | 1                      |                                |
| hCMV-FL-pol [E3+]   | Pool | 0.90*, 90% |   | 56                     | 4.2  | 4.7                            | 168                    | 6.5                            |
|                     | 1    |            | 1.18, 88%                                       |                        |  |                                |                        |                                |
|                     | 2    |            | 1.04, 80%                                       |                        |  |                                |                        |                                |

MRKAd5nef and MRKAd5pol Viral Production Kinetics - A timecourse experiment was carried out in roller bottles to determine if the viral production kinetics of the MRKAd5pol and MRKAd5nef vectors were similar to those of MRKAd5gag. PER.C6® cells in roller bottle cultures were infected at an MOI of 280 vp/cells with P5 MRKAd5pol, P5 MRKAd5nef and P7 MRKAd5gag; for each adenovector, two infected bottles were sampled at 24, 36, 48, and 60 hours post infection. In addition, two bottles were left unsampled until 48 hpi when they were harvested under the Phase I process conditions. The anion-exchange HPLC viral particle concentrations of the freeze-thaw recovered cell associated virus at the 24, 36,

48, and 60 hpi timepoints are shown in Figure 29A-B. The QPA titers show a similar trend (data not shown).

Comparison of hCMV- and mCMV-FL-nef - As the titers obtained with the MRKAd5nef construct (hCMV-FL-nef) were lower than those obtained with MRKAd5gag or MRKAd5pol, a viral productivity comparison experiment was performed with mCMV-FL-nef. For each of the two adenovectors (hCMV- and mCMV-FL-nef), two roller bottles were infected at an MOI of 280 vp/cell with passage five clarified lysate. The macroscopic and microscopic observations of the four roller bottles were identical at the time of harvest. Analysis of the clarified lysate produced indicated a higher viral particle concentration in the bottles infected with mCMV-FL-nef, as shown in Table 19. It is stipulated that the higher productivity with mCMV promoter driven nef vector is due to lower nef expression levels in PER.C6® cells- experiments are underway at V&CB to measure nef expression levels.

15

Table 19. Passage Six Viral Productivity Comparison of hCMV- and mCMV-FL-nef

|             | [    | Xv (10 <sup>6</sup> cells/m | l), Viability (%) | Cell Passage | AEX Titer                      | Titer                   | Amplification | Triton Lysis Titer             |
|-------------|------|-----------------------------|-------------------|--------------|--------------------------------|-------------------------|---------------|--------------------------------|
| •           |      | Infection                   | Harvest           | Number       | 10 <sup>10</sup> vp/ml culture | 10 <sup>4</sup> vp/cell | Ratio         | 10 <sup>10</sup> vp/ml culture |
| hCMV-FL-nef | Pool | 1.11,91%                    |                   | 60           | 1.5                            | 1.4                     | 50            | 2.8                            |
| (MRKAd5nef) | i    |                             | 1.23,75%          |              |                                |                         |               |                                |
|             | 2    |                             | 1.34, 74%         |              | •                              |                         |               |                                |
| mCMV-FL-nef | Pool | 1.11, 91%                   | 1                 | 60           | 2.3                            | 2.1                     | 75            | 4.6                            |
|             | 1    |                             | 1.49, 84%         |              |                                |                         |               |                                |
|             | 2    |                             | 1.18, 77%         | 1            |                                |                         |               |                                |

20

25

30

#### **EXAMPLE 27**

# Characterization and Large Scale Production of MRKAd5nef Virus in Bioreactors

Materials and Methods - The experiment of the present example was run twice under the following conditions: 36.5°C, DO 30%, pH 7.30, 150rpm agitation rate, no sparging, Life Technologies (Gibco, Invitrogen) 293 SFM II (with 6mM L-glutamine), 0.5M NaOH as base for pH control. During the first run (B20010115), two 10L stirred vessel bioreactors were inoculated with PER.C6® cells at a concentration of 0.2x106 cells/ml. Cells were grown until they reached a cell concentration of approximately 1x106 cells/ml. The cells were infected with uncloned MRKAd5nef (G2A,LLAA) at a MOI of 280 virus particles (vp)/cell. For the second batch (B20010202), the same procedure as the first run was used, except the cells

were infected with cloned MRAd5nef. During both runs, the bioreactors were harvested 48 hours post-infection. Samples were taken and virus concentrations were determined from whole broth (with triton lysis), supernatant, and cell pellets (3 X freeze/thaw) with the AEX and QPA assays. Metabolites were measured with BioProfile 250 throughout the process.

Table 20: Experimental Conditions

| Temperature | 36.5 ℃  |  |
|-------------|---------|--|
| DO          | 30%     |  |
| PH          | 7.30    |  |
| Agitation   | 150 rpm |  |
| Sparging    | None    |  |

Table 21: Virus source used for experiments.

10

| Run | Batch ID    | Cloned/Uncloned<br>MRKAd5nef | MOI<br>(vp/cells) |
|-----|-------------|------------------------------|-------------------|
| #1  | B20010115-1 | Uncloned                     | 280               |
| _   | B20010115-2 | Uncloned                     | 280               |
| #2  | B20010202-1 | Cloned                       | 280               |
| l   | B20010202-2 | Cloned                       | 280               |

Results - Table 22 and 23 show an the ability to scale up production of MRKAd5nef by growth in a bioreactor.

15

Table 22: Virus Concentration as measured by the AEX assay

| Run | Batch ID    | Cloned/Uncloned | V           | irus Concentration @ | ⊉ 48hpi (1x | 10 <sup>13</sup> vp/L) |
|-----|-------------|-----------------|-------------|----------------------|-------------|------------------------|
|     |             | MRKAd5nef       | Supernatant | Clarified Lysate     | Total       | Triton Lysate          |
| #1  | B20010115-1 | Uncloned        | 0.72        | 3.26                 | 3.98        | 5.76                   |
|     | B20010115-2 | Uncloned        | 0.38        | 1.67                 | 2.05        | 2.46                   |
| #2  | B20010202-1 | Cloned          | 0.80        | 6.00                 | 6.80        | 8.88                   |
|     | B20010202-2 | Cloned          | 0.50        | 6.00                 | 6.50        | 8.47                   |

Table 23: Virus Titers as measured by the QPA assay

| Run | Batch ID    | Cloned/Uncloned |       | Virus Concentration @ 48hpi (1x10 <sup>11</sup> IU/L) |           |       |        |  |  |  |
|-----|-------------|-----------------|-------|---|-----------|-------|--------|--|--|--|
| 1   |             | MRKAd5nef       | Whole | Supernatant   | Clarified | Total | Triton |  |  |  |
|     |             |                 | Broth |   | Lysate    |       | Lysate |  |  |  |
| #1  | B20010115-1 | Uncloned        | 0.13  | 1.12  | 1.76      | 2.88  | 11.28  |  |  |  |
| Į   | B20010115-2 | Uncloned        | 0.14  | 0.73  | 1.54      | 2.27  | 5.86   |  |  |  |
| #2  | B20010202-1 | Cloned          | 0.14  | 0.97  | 1.62      | 2.69  | 11.89  |  |  |  |
|     | B20010202-2 | Cloned          | 0.14  | 1.17  | 1.70      | 2.97  | 12.47  |  |  |  |

20

The present invention is not to be limited in scope by the specific embodiments described herein. Indeed, various modifications of the invention in addition to those described herein will become apparent to those skilled in the art

from the foregoing description. Such modifications are intended to fall within the scope of the appended claims.

#### **EXAMPLE 28**

## MRKAd5HIV-1gag Boosting of DNA-Primed Animals

5

10

15

20

25

30

Groups of 3-5 rhesus macaques were immunized with (a) 5 mgs of V1Jns-Flgag (pVIJnsCMV(no intron)-FL-gag-bGHpA), (b) 5 mgs of V1Jns-Flgag formulated with 45 mgs of a non-ionic block copolymer CRL1005, or (c) 5 mgs of V1Jns-Flgag formulated with 7.5 mgs of CRL1005 and 0.6 mM benzalkonium chloride at weeks 0, 4, and 8. All animals received a single dose of 10e7 viral particles (vp) of the MRKAd5HIV-1gag at week 26. Note: 10e7 is too low to prime or boost effectively when used as a single modality (dose is selected to mimic preexposure to adenovirus); see Figure 32.

Blood samples were collected from all animals at several time points and peripheral blood mononuclear cells (PBMCs) were prepared using standard Ficoll method. The PBMCs were counted and analyzed for gamma-interferon secretion using the ELISpot assay (Table 24). For each monkey, the PBMCs were incubated overnight either in the absence (medium) or presence of a pool (called "gag H") of 50 20-aa long peptides that encompass the entire HIV-1 gag sequence.

The results indicate that MRKAd5HIV-1gag was very effective in boosting the T cell immune responses in these monkeys. At week 28 or 2 weeks after the viral boost, the number of gag-specific T cells per million PBMCs increased 2-48 fold compared to the levels observed at week 24 or 2 weeks prior to the boost.

The PBMCs were also analyzed by intracellular gamma-interferon staining prior to (at week 10) and after the MRKAd5gag boost (at week 30). The results for select animals are shown on Figure 31. The results indicate that (a) immunization with DNA/adjuvant formulation elicited T cell responses which can either be balanced, CD4<sup>+</sup>-biased or CD8<sup>+</sup>-biased, and (b) boosting with the MRKAd5gag construct produced in all cases a strongly CD8<sup>+</sup>-biased response. These results suggest that boosting with MRKAd5HIV-1gag construct is able to improve the levels of antigen-specific CD8<sup>+</sup> T cells.

Table 24. Boosting of DNA/Adjuvant-Primed Rhesus Monkeys with MRKAd5gsg Number of SFCmillion PBMCs

| ā     | Priming                      | Boost          | Wonk#  | 2      |       | ۳      | Ya4          | J≖E    | و    | T=10   | 0     | T=17       | _     | T=24     | 4     | T=28   | 88    | T=30     |       |
|-------|------------------------------|----------------|--------|--------|-------|--------|--------------|--------|------|--------|-------|------------|-------|----------|-------|--------|-------|----------|-------|
|       | FacU, 4, 8 WICE              | Ta26 WKs       |        | Medium | H ded | Medium | <b>D80</b> H | Medium | _    | Medium | H CBD | Wedlum     | -     | Medium   | H DBD | Medium | H DED | Medium   | H OBD |
| _     | DNA/6 mgs                    | MRKAd5gag(E3+) | CBSH   | ¥      | ΑĀ    | 3      | æ            | 1      | F    | -      | 22    |            | 115   | 6        | 22    | 2      | 926   | ٥        | 318   |
|       | PBS                          | 10^7 vp        | š      | 0      | 0     | 0      | 5            | 0      | 46   | ٥      | 99    | 0          | 76    | 0        | 32    | 6      | 1705  | _        | 755   |
|       | (D101)                       |                | AW3G   | s)     | =     | 0      | 98           |        | 19   | 6      | 48    | 2          | 68    | <b>5</b> | 89    | 2      | 686   | •        | 382   |
| ~     | DNA/6mgs +                   | MRKAd5can(E3+) | CCTC   | -      | -     | -      | 8            |        | =    | 4      | 27.0  | †          | , occ |          | 333   | 1,     | 2     | ţ        |       |
|       | CRL1005/45mgs .              | 10.7 va        | CCIK   | . 4    | -     | _      | 3 5          | -      | 25.4 |        | 2 5   |            | 68    |          | 3 5   | , (    |       | <u>.</u> | 9 5   |
|       | •                            |                | AWB    | . 6    | 60    | -      | 2            | . **   | 7    | . 4    | 4     | . «        | 2 5   | - L      | ž #   | > =    | 2 2   | - «      | 24.5  |
| _     |                              |                | CBSF   | ž      | ž     | 0      | 8            | 0      | 88   | -      | 230   | - <u>6</u> | 374   |          | 792   | : 00   | 1548  | 8        | 1734  |
|       |                              |                | AKBB   | 6      | 5     | 4      | 8            | -      | 119  | •      | 439   | •          | 425   | 0        | 318   | 4      | 1229  | 10       | 354   |
| 6     | DNA/5 mgs+                   | MRKAd5gag(E3+) | AW20   | 9      | 4     | -      | 59           | la.    | 264  | 9      | 425   | 9          | 55    | ۰        | 205   | ē      | FRF   | •        | ş     |
|       | CRL1005/7.5 mgs + 0.8 mM BAK | 10v7 vp        | CAB    | -      | 0     | 6      | 121          | _      | 55   | -      | 23    | LO.        | 8     |          | 193   | =      | 1384  | £        | 878   |
|       |                              |                | 888    | 8      | 9     | •      | 9            | 60     | 119  | 0      | 274   | 9          | 282   | _        | 28    | 0      | 83    | -        | 828   |
|       |                              |                | CB6W   | 4      | 6     | 0      | 88           | -      | 16   | Ö      | 139   | 0          | 2     | -        | 8     | ı,     | 643   | _        | 349   |
|       |                              |                | CB7D   | -      | 0     | •      | 136          | 0      | 316  | -      | 609   | us         | 626   | -        | 769   | •      | 2278  | •        | 豆     |
| 4     | none                         | None           | 980201 | 5      | 0     | 0      | 0            | -      | ŀ    | ŀ      | °     | ŀ          | -     | -        | ~     | 6      | -     |          | c     |
| not a | ot available                 |                |        |        |       |        |              |        | 1    |        |       |            |       |          | 1     | ,      | •     |          | ,     |

#### **EXAMPLE 29**

## Construction of gagpol fusion for MRKAd5gagpol fusion constructs

The open reading frames for the codon-optimized HIV-1 gag gene was fused directly to the open reading frame of the IA pol gene (consisting of RT, RNAseH and integrase domains) by stepwise PCR. Because the gene (SEQ ID NO: 38) does not include the protease gene and the frameshift sequence, it encodes a single polypeptide of the combined size of p55, RT, RNAse H and integrase (1350 amino acids; SEQ ID NO: 39).

5

10

15

20

25

30

35

The fragment that extends from the BstEII site within the gag gene to the last non-stop codon was ligated via PCR to a fragment that extends from the start codon of the IApol to a unique BamHI site. This fragment was digested with BstEII and BamHI. Construction of gag-IApol fusion was achieved via three-fragment ligation involving the PstI-BstEII gag digestion fragment, the BstEII/BamHI digested PCR product and long PstI/BamHI V1R-FLpol backbone fragment.

The MRKAd5-gagpol adenovirus vector was constructed using the BglII fragment of the V1R-gagpol containing the entire ORF of gag-IApol fusion gene.

#### **EXAMPLE 30**

Immunogenicity Studies in Non-Human Primates

Cohorts of three (3) macaques were immunized with 10e8 or 10e10 viral particles (vp) of one of the following MRKAd5 HIV-1 vaccines: (1) MRKAd5gag; (2) MRKAd5pol; (3) MRKAd5nef; (4) a mixture containing equal amounts of MRKAd5gag, MRKAd5pol, and MRKAd5nef, or (5) a mixture of equal amounts of MRKAd5gagpol and MRKAd5nef. The vaccines were administered at weeks 0 and 4.

The T cell responses against each of the HIV-1 antigens were assayed by IFN-gamma ELISpot assay using pools of 20-aa peptides that encompass the entire protein sequence of each antigen. The results (Table 25) are expressed as the number of spot-forming cells (sfc) per million peripheral blood mononuclear cells (PBMC) that respond to each of the peptide pools.

Results indicate the following observations: (1) each of the single gene constructs (MRKAd5gag, MRKAd5pol, or MRKAd5nef) is able to elicit high levels of antigen-specific T cells in monkeys; (2) the single-gene MRKAd5 constructs can be mixed as a multi-cocktail formulation capable of eliciting very broad T cell responses against gag, pol, and nef; (3) the MRKAd5 vector expressing the fusion

protein of gag plus IA pol is capable of inducing strong T cell responses to both gag and pol.

Table 25. Evaluation of Mixtures of MRKAd5 vectors expressing humanized

5 HIV-1 gag, pol, gagpol, nef in rhesus macaques

10

| Grp# | Vaccine                       | Monk#  |      |       | T=6 wks |         |      |
|------|-------------------------------|--------|------|-------|---------|---------|------|
|      | T=0, 4 wks                    | l      | Mock | Gag H | Pol - 1 | Pol - 2 | Nef  |
| 1    | MRKAd5 gag                    | CB9V   | 0    | 15    | -       | -       | -    |
|      | 10^10 vp                      | CD19   | 0.   | 374   |         | -       | -    |
|      |                               | 109H   | 1    | 843   | -       | -       | -    |
| 2    | MRKAd5 gag                    | 99D130 | 1    | 948   | -       |         | -    |
|      | 10^8 vp                       | W277   | 16   | 324   | -       | -       | -    |
|      |                               | 143H   | . 4  | 595   | -       | -       | -    |
| 3    | MRKAd5 pol                    | CC1X   | 4    | -     | 46      | 256     | -    |
|      | 10^10 vp                      | AW3W   | 3    | -     | 463     | 550     | -    |
| 1    |                               | AV43   | 6    | -     | 95      | 1333    | -    |
| 4    | MRKAd5 pol                    | AW38   | 1    | -     | 19      | 30      | -    |
| ļ    | 10^8 vp                       | CC8K   | 0    | -     | 50      | 995     | -    |
|      |                               | CC21   | 1    | -     | 33      | · 436   | -    |
| 5    | MRKAd5 nef                    | 076Q   | 9    | -     | -       | -       | 1204 |
|      | 10^10 vp                      | 091Q   | 4    |       | -       | -       | 85   |
|      |                               | 083Q   | 0    | -     | -       | -       | 176  |
| 6    | MRKAd5 nef                    | 00C029 | 1    | -     | -       | -       | 114  |
| İ    | 10^8 vp                       | 98D022 | 6    | -     | -       | -       | 170  |
|      |                               | 98D160 | 3    | -     | -       | -       | 198  |
| 7    | MRKAd5gag+MRKAd5pol+MRKAd5nef | 99D251 | 3    | 206   | 15      | 193     | 120  |
|      | 10^10 vp each                 | 05H    | 3    | 135   | 21      | 9       | 638  |
|      | ·                             | 00C016 | 3    | 26    | 4       | 51      | 23   |
| 8    | MRKAd5gag+MRKAd5pol+MRKAd5nef | 99D215 | 1    | 171   | 18      | 193     | 240  |
| l    | 10^8 vp each                  | 81H    | 5    | 73    | 6       | 14      | 243  |
|      |                               | -12H   | 8    | 1140  | 115     | 811     | 719  |
| 9    | MRKAd5gagpol +MRKAd5 nef      | 99D211 | 0    | 83    | 56      | 838     | 725  |
|      | 10^10 vp each                 | 22H    | 4    | 385   | 119     | 1194    | 1915 |
|      |                               | 61H    | 4    | 343   | 11      | 765     | 853  |
| 10   | MRKAd5gagpol +MRKAd5 nef      | 34H    | 3    | 78    | 19      | 5       | 75   |
| j    | 10^8 vp each                  | 48H    | 1    | 65    | 105     | 46      | 43   |
| İ    |                               | 70H    | 5    | 158   | 15      | 220     | 191  |

Indicated are numbers of spot-forming cells per million PBMCS against the peptide pools. Mock, no peptides; gag H, fifty 20-aa peptides encompassing p55 sequence; pol-1, 20-aa peptides representing N-terminal half of IA pol; pol-2, 20-aa peptides representing the carboxy-terminal half of IA pol; nef, 20-aa peptides encompassing the entire wild-type nef sequence. Responses to the antigens prior to the first immunization did not exceed 40 sfc/10^6 PBMC.

#### WHAT IS CLAIMED IS

10

1. A recombinant adenoviral vaccine vector at least partially deleted in

- 5 E1 and devoid of E1 activity, comprising:
  - a) an adenovirus cis-acting packaging region corresponding to from about base pair 1 to between from about base pair 400 to about base pair 458 of a wildtype adenovirus genome; and
  - b) a gene encoding an HIV protein or immunologically relevant modification thereof.
  - A vector in accordance with claim 1 comprising a packaging region corresponding to from about base pair 1 to about base pair 450 of a wildtype adenovirus genome.
- 3. A vector in accordance with claim 1 further comprising nucleotides
   15 corresponding to between from about base pair 3511 to about 3524 to about base pair
   5798 of a wildtype adenovirus genome.
  - 4. A vector in accordance with claim 3 comprising base pairs corresponding to 1-450 and 3511-5798 of a wildtype adenovirus genome.
- 5. A vector in accordance with claim 4 which is deleted of base pairs451-3510.
  - 6. A vector in accordance with claim 1 which is at least partially deleted in E3.
  - 7. A vector in accordance with claim 6 wherein the E3 deleted region is from base pairs 28,133-30,818.

8. A vector in accordance with claim 1 wherein the gene encoding the HIV protein or modification thereof comprises codons optimized for expression in a human.

- 9. A vector in accordance with claim 1 wherein the vector comprises a5 gene expression cassette comprising:
  - a) a nucleic acid encoding a protein;
  - b) a heterologous promoter operatively linked to the nucleic acid encoding the protein; and
    - (c) a transcription termination sequence.
- 10. A vector in accordance with claim 9 wherein the gene expression cassette is inserted into the E1 region.
  - 11. An adenoviral vector in accordance with claim 9 wherein the gene expression cassette is in an E1 parallel orientation
- 12. An adenoviral vector in accordance with claim 9 wherein the geneexpression cassette is in an E1 antiparallel orientation.
  - 13. An adenoviral vector in accordance with claim 9 wherein the promoter is a cytomegalovirus promoter devoid of intronic sequences.
  - 14. An adenoviral vector in accordance with claim 13 wherein the promoter is an immediate early human cytomegalovirus promoter.
- 20 15. An adenoviral vector in accordance with claim 9 wherein the promoter is a murine cytomegalovirus promoter.
  - 16. An adenoviral vector in accordance with claim 9 wherein the transcription termination sequence is a bovine growth hormone polyadenylation and transcription termination sequence.

17. An adenoviral vector in accordance with claim 9 wherein the transcription termination sequence is a synthetic polyadenylation signal (SPA).

- 18. A cell comprising the adenoviral vector of claim 1.
- 19. Recombinant, replication-defective adenovirus particles harvested
   and purified subsequent to transfection of the adenoviral vector of claim 1 into a cell
   line which expresses adenovirus E1 protein at complementing levels.
  - 20. An HIV vaccine composition comprising purified adenovirus particles of claim 19.
- 21. An HIV vaccine composition of claim 20 which comprises aphysiologically acceptable carrier.
  - 22. A method of producing recombinant, replication defective adenovirus particles containing the adenoviral genome of the adenoviral vector of claim 1 which comprises introducing the adenoviral vector into a host cell which expresses adenoviral E1 protein, and harvesting the resultant recombinant, replication-defective adenovirus.
  - 23. A method according to claim 22 wherein the cell is a PER.C6<sup>®</sup> cell.

15

- 24. A method of generating a cellular-mediated immune response against HIV in an individual comprising administering to the individual a vaccine of claim 21.
  - 25. A method according to claim 24 which further comprises administration to the individual a DNA plasmid vaccine, optionally administered with a biologically effective adjuvant, protein or other agent capable of increasing the immune response.

26. A method according to claim 25 wherein the DNA plasmid vaccine is administered to the individual prior to administration of an adenovirus vaccine.

- 27. A method according to claim 24 wherein the adenovirus vaccine is5 preceded by an adenovirus vaccine of a different serotype.
  - 28. A method according to claim 24 which comprises administering and readministering the adenovirus vaccine vector to the individual.
  - 29. An adenoviral vector in accordance with claim 1 wherein the HIV protein is HIV gag or an immunologically relevant modification thereof.
- 30. An adenoviral vector in accordance with claim 9 wherein the gene expression cassette comprises an open reading frame encoding an HIV gag protein or immunologically relevant modification thereof.
  - 31. A recombinant adenoviral vaccine vector at least partially deleted in E1 and devoid of E1 activity, comprising:
- a) an adenovirus *cis*-acting packaging region corresponding to from about base pair 1 to about base pair 450 and from about 3511 to about 5798 of a wildtype adenovirus genome, and deleted for base pairs corresponding to from about base pair 451 to from about base pair 3510 of a wildtype adenovirus genome; and
  - b) a gene expression cassette comprising

- i) SEQ ID NO: 29;
- ii) a heterologous promoter operatively linked to i); and
- iii) a transcription termination sequence.

32. An adenoviral vector in accordance with claim 31 wherein the gene expression cassette is in an E1 parallel orientation.

- 33 An adenoviral vector in accordance with claim 31 wherein the gene expression cassette is in an E1 antiparallel orientation.
- 34. An adenoviral vector in accordance with claim 31 wherein the promoter is a cytomegalovirus promoter devoid of intronic sequences.

5

10

15

- 35. An adenoviral vector in accordance with claim 31 wherein the transcription termination sequence is a bovine growth hormone polyadenylation and transcription termination sequence.
- 36. An adenoviral vector in accordance with claim 31 which is at least partially deleted in E3.
  - 37. A cell comprising the adenoviral vector of claim 30.
- 38. Recombinant, replication-defective adenovirus particles harvested and purified subsequent to transfection of the adenoviral vector of claim 30 into a cell line which expresses adenovirus E1 protein at complementing levels.
- 39. An HTV vaccine composition comprising purified adenovirus particles of claim 38.
- 40. An HIV vaccine composition of claim 39 which comprises a physiologically acceptable carrier.
- 41. A method of producing recombinant, replication defective adenovirus particles containing the adenoviral genome of the adenoviral vector of claim 30 which comprises introducing the adenoviral vector into a host cell which expresses adenoviral E1 protein, and harvesting the resultant recombinant, replication-defective adenovirus.

42. A method according to claim 41 wherein the cell is a PER.C6® cell.

43. A method of generating a cellular-mediated immune response against HIV in an individual comprising administering to the individual a vaccine of claim 21.

- 44. A method according to claim 43 which further comprises administration to the individual a DNA plasmid vaccine, optionally administered with a biologically effective adjuvant, protein or other agent capable of increasing the immune response.
- 45. A method according to claim 44 wherein the DNA plasmid vaccine is administered to the individual prior to administration of an adenovirus vaccine.
  - 46. A method according to claim 43 wherein the adenovirus vaccine is preceded by an adenovirus vaccine of a different serotype.
- 47. A method according to claim 43 which comprises administering and readministering the adenovirus vaccine vector to the individual.
  - 48. An adenoviral vector in accordance with claim 1 wherein the HIV protein is HIV pol or an immunologically relevant modification thereof.
- 49. An adenoviral vector in accordance with claim 9 wherein the gene
   20 expression cassette comprises an open reading frame encoding an HIV pol protein or immunologically relevant modification thereof.
  - 50. A recombinant adenoviral vaccine vector at least partially deleted in E1 and devoid of E1 activity, comprising:

a) an adenovirus *cis*-acting packaging region corresponding to from about base pair 1 to about base pair 450 and from about 3511 to about 5798 of a wildtype adenovirus genome, and deleted for base pairs corresponding to from about base pair 451 to from about base pair 3510 of a wildtype adenovirus genome; and

5

10

15

- b) a gene expression cassette comprising
  - i) a nucleotide sequence selected the group consisting of SEQ ID NO: 1, SEQ ID NO: 5 and SEQ ID NO: 7;
  - ii) a heterologous promoter operatively linked to i); and
- iii) a transcription termination sequence.
- 51. An adenoviral vector in accordance with claim 50 wherein the gene expression cassette is in an E1 parallel orientation.
- 52. An adenoviral vector in accordance with claim 50 wherein the gene expression cassette is in an E1 antiparallel orientation.
- 53. An adenoviral vector in accordance with claim 50 wherein the promoter is a cytomegalovirus promoter devoid of intronic sequences.
- 54. An adenoviral vector in accordance with claim 50 wherein the transcription termination sequence is a bovine growth hormone polyadenylation and transcription termination sequence.
- 55. An adenoviral vector in accordance with claim 50 which is at least partially deleted in E3.
  - 56. A cell comprising the adenoviral vector of claim 49.

57. Recombinant, replication-defective adenovirus particles harvested and purified subsequent to transfection of the adenoviral vector of claim 49 into a cell line which expresses adenovirus E1 protein at complementing levels.

58. An HTV vaccine composition comprising purified adenovirus particles of claim 57.

5

10

- 59. An HIV vaccine composition of claim 58 which comprises a physiologically acceptable carrier.
- 60. A method of producing recombinant, replication defective adenovirus particles containing the adenoviral genome of the adenoviral vector of claim 49 which comprises introducing the adenoviral vector into a host cell which expresses adenoviral E1 protein, and harvesting the resultant recombinant, replication-defective adenovirus.
- 61. A method according to claim 60 wherein the cell is a PER.C6® cell.
- 62. A method of generating a cellular-mediated immune response against HIV in an individual comprising administering to the individual a vaccine of claim 59.
  - 63. A method according to claim 62 which further comprises administration to the individual a DNA plasmid vaccine, optionally administered with a biologically effective adjuvant, protein or other agent capable of increasing the immune response.
  - 64. A method according to claim 63 wherein the DNA plasmid vaccine is administered to the individual prior to administration of an adenovirus vaccine.

65. A method according to claim 62 wherein the adenovirus vaccine is preceded by an adenovirus vaccine of a different serotype.

- 66. A method according to claim 62 which comprises administering and readministering the adenovirus vaccine vector to the individual.
- 67. An adenoviral vector in accordance with claim 1 wherein the HIV protein is HIV nef or an immunologically relevant modification thereof.

5

10

15

20

- 68. An adenoviral vector in accordance with claim 9 wherein the gene expression cassette comprises an open reading frame encoding an HIV nef protein or immunologically relevant modification thereof.
- 69. A recombinant adenoviral vaccine vector at least partially deleted in E1 and devoid of E1 activity, comprising:
  - a) an adenovirus *cis*-acting packaging region corresponding to from about base pair 1 to about base pair 450 and from about 3511 to about 5798 of a wildtype adenovirus genome, and deleted for base pairs corresponding to from about base pair 451 to from about base pair 3510 of a wildtype adenovirus genome; and
  - b) a gene expression cassette comprising
    - a nucleotide sequence selected the group consisting of SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13 and SEQ ID NO: 15;
    - ii) a heterologous promoter operatively linked to i); and
    - iii) a transcription termination sequence.
- 70. An adenoviral vector in accordance with claim 69 wherein the gene expression cassette is in an E1 parallel orientation.

71. An adenoviral vector in accordance with claim 69 wherein the gene expression cassette is in an E1 antiparallel orientation.

- 72. An adenoviral vector in accordance with claim 69 wherein the promoter is a cytomegalovirus promoter devoid of intronic sequences.
- 5 73. An adenoviral vector in accordance with claim 69 wherein the transcription termination sequence is a bovine growth hormone polyadenylation and transcription termination sequence.
  - 74. An adenoviral vector in accordance with claim 69 which is at least partially deleted in E3.
    - 75. A cell comprising the adenoviral vector of claim 68.

10

15

20

- 76. Recombinant, replication-defective adenovirus particles harvested and purified subsequent to transfection of the adenoviral vector of claim 68 into a cell line which expresses adenovirus E1 protein at complementing levels.
- 77. An HIV vaccine composition comprising purified adenovirus particles of claim 76.
  - 78. An HIV vaccine composition of claim 77 which comprises a physiologically acceptable carrier.
- 79. A method of producing recombinant, replication defective adenovirus particles containing the adenoviral genome of the adenoviral vector of claim 68 which comprises introducing the adenoviral vector into a host cell which expresses adenoviral E1 protein, and harvesting the resultant recombinant, replication-defective adenovirus.
- 80. A method according to claim 79 wherein the cell is a PER.C6® cell.

81. A method of generating a cellular-mediated immune response against HIV in an individual comprising administering to the individual a vaccine of claim 78.

- 82. A method according to claim 81 which further comprises

  administration to the individual a DNA plasmid vaccine, optionally administered with a biologically effective adjuvant, protein or other agent capable of increasing the immune response.
  - 83. A method according to claim 82 wherein the DNA plasmid vaccine is administered to the individual prior to administration of an adenovirus vaccine.

10

15

20

- 84. A method according to claim 81 wherein the adenovirus vaccine is preceded by an adenovirus vaccine of a different serotype.
- 85. A method according to claim 81 which comprises administering and readministering the adenovirus vaccine vector to the individual.
- 86. A multivalent adenovirus vaccine composition comprising recombinant, replication-defective adenovirus particles, wherein the adenovirus particles are harvested and purified from a cell line expressing adenovirus E1 protein, and wherein the particles are harvested subsequent to transfection of the cells with an adenoviral vector or vectors in accordance with claim 9; said vector(s) comprising a gene expression cassette or cassettes comprising nucleotide sequences encoding HIV proteins selected from the group consisting of:
  - gag, pol, and nef, expressed independently from three individual vectors;

WO 02/22080

b) gag, pol, and nef, expressed independently from one vector with the encoding nucleic acid sequences operatively linked to distinct promoters and transcription termination sequences; c) gag, pol, and nef, expressed via two vectors, one expressing a pol-. 5 nef fusion, and another expressing gag; d) gag, pol, and nef, expressed via two vectors, one expressing a gagpol fusion and another expressing nef; e) gag, pol and nef, expressed via two vectors, one expressing a nefgag fusion and another expressing pol; 10 gag, pol, and nef, expressed via one vector expressing a gag-polnef fusion; gag and pol, expressed independently from two individual vectors; h) gag and pol, expressed independently from one vector with the encoding nucleic acid sequences operatively linked to distinct 15 promoters and transcription termination sequences; pol and nef, expressed independently from two individual vectors; pol and nef, expressed independently from one vector with the encoding nucleic acid sequences operatively linked to distinct promoters and transcription termination sequences; 20 k) nef and gag, expressed independently from two individual vectors; nef and gag, expressed independently from one vector with the encoding nucleic acid sequences operatively linked to distinct promoters and transcription termination sequences; m) gag and pol, expressed via one vector expressing a gag-pol fusion;

n) pol and nef, expressed via one vector expressing a pol-nef fusion; and

- o) nef and gag, expressed via one vector expressing a nef-gag fusion.
- 87. A multivalent adenovirus vaccine composition in accordance with claim 86 wherein the gag-pol fusion consists of SEQ ID NO: 39.
  - 88. A multivalent adenovirus vaccine composition in accordance with claim 86 wherein the fused sequences have the encoding nucleic acid sequences operatively linked to distinct promoters and transcription termination sequences.
- 89. A multivalent adenovirus vaccine composition in accordance with

  10 claim 86 wherein the fused sequences have the encoding nucleic acid sequences

  operatively linked to a single promoter; and the encoding nucleic acid sequences

  operatively linked by an internal ribosome entry sequence ("TRES").

#### Original Adenovector Construct:

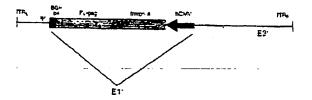


Figure 1: Original HIV-1 gag adenovector.

### Sequence of the open reading frame for FL-gag (human codon optimized)

atgggtgctagggcttctgtgctgtctggtggtgagctggacaagtgggagaagatcaggctgaggcctggtgg caagaagaagtacaagctaaagcacattgtgtgggcctccagggagctggagaggtttgctgtgaaccctggc agctgaggiccctgtacaacacagtggctaccctgtactgtgtgcaccagaagattgatgtgaaggacaccaag gaggccctggagaagattgaggaggagcagaacaagtccaagaagaaggcccagcaggctgctgctggc acaggcaactccagccaggtgtcccagaactaccccattgtgcagaacctccagggccagatggtgcaccag gccatctcccccggaccctgaatgcctgggtgaaggtggtggaggagaaggccttctcccctgaggtgatccc catgitictetgecetgictgagggtgecacceccaggacetgaacaccatgetgaacacagtggggggecate aggetgecatgeagatgetgaaggagaceateaatgaggaggetgetgagtgggacaggetgeateetgtge acgctggccccattgcccccggccagatgagggagcccaggggctctgacattgctggcaccacctccaccct ccaggagcagattggctggatgaccaaccaccccccatccctgtggggggaaatctacaagaggtggatcat cccttcagggactatgtggacaggttctacaagaccctgagggctgagcaggcctcccaggaggtgaagaact ggatgacagagaccctgctggtgcagaatgccaaccctgactgcaagaccatcctgaaggccctgggccctg gctgaggccatgtcccaggtgaccaactccgccaccatcatgatgcagaggggcaacttcaggaaccagag gaagacagtgaagtgcttcaactgtggcaaggtgggccacattgccaagaactgtagggcccccaggaaga ggcaaaatctggccctcccacaagggcaggcctggcaacttcctccagtccaggcctgagcccacagcccct agetglaccccetggcctccctgaggtccctgtttggcaacgacccctcctcccagtaaaataaagcccgggca gat (SEQ ID NO: 29)

Figure 2

#### Old Transgene:



#### New Transgenes:

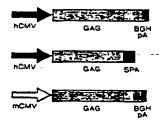


Figure 3: Diagrammatic representation of the original HIV-1 gag transgene and the series of new transgene constructions.

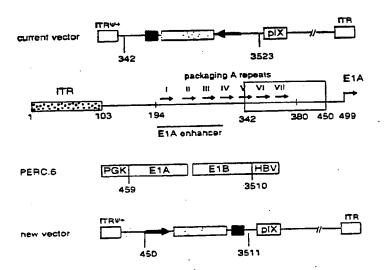


Figure 4: Modifications made to the current adenovector backbone in the generation of the new vector.

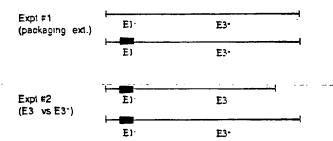


Figure 5: Virus mixing experiments to determine the effects of the addition made to the packaging signal region (Expt #1) and analysis of the effects of the E3 gene on viral growth (Expt. #2). The red bars denote the region of modifications made to the E1 deletion.



Figure 6: Autoradiograph of viral DNA analysis following viral mixing experiments (expts. #1 and #2) as detailed in the text.

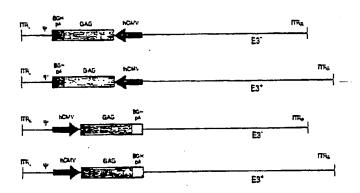


Figure 7A: hCMV-FLgag-bGHpA adenovectors constructed within the "MRK" backbone. E1 parallel and E1 antiparallel transgene orientation within the E3- and E3+ backbones were constructed.

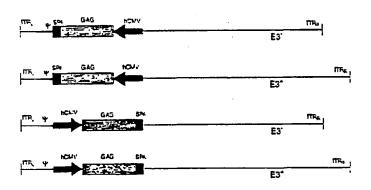


Figure 7B: hCMV-FLgag-SPA adenovectors constructed within the "MRK" backbone. E1 parallel and E1 antiparallel transgene orientation within the E3- and E3+ backbones were constructed.

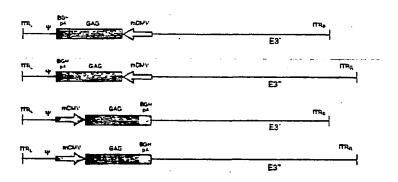


Figure 7C: mCMV-FLgag-bGHpA adenovectors constructed within the "MRK" backbone. E1 parallel and E1 antiparallel transgene orientation within the E3- and E3+ backbones were constructed.

### Plasmid mixing expt: (orientation)

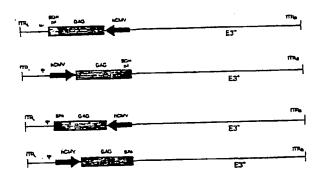


Figure 8A: Effect of transgene orientation

#### Plasmid Mixing expt: (poly A signal)

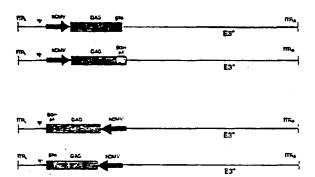


Figure 8B: Effect of polyadenylation signal



Figure 9: Viral DNA from the four Adgag candidates at P5, following EstE11 digestion.

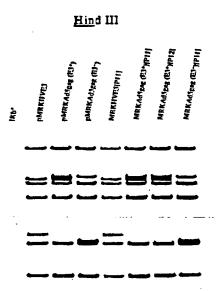


Figure 10: Viral DNA analysis of passage 11 and/or 12 of MRKHVE3, MRKAd5gag and MRKAd5gag(E3-).

PCT/US01/28861

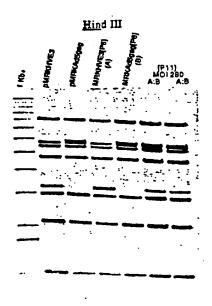


Figure 11: Viral DNA analysis (*Hin*dIII digestion) of passage 6 MRKHVE3 and MRKAd5gag used to initiate the viral competition study. Last two lanes are passage 11 analysis of duplicate passages of the competition study (each virus at MOI 280 vp).

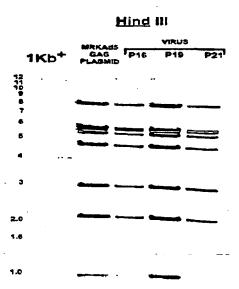
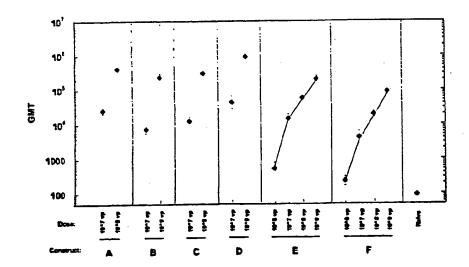


Figure 12: Viral DNA analysis by *Hin*dIII digestion on high passage numbers for MRKAd5gag in serum containing media with collections made at specified times. The first lane shows the 1 Kb DNA size marker. The other lanes represent pre-plasmid control (digested with Pac1 and *Hin*dIII), and MRKAd5gag virus continually passaged to P16, P19 and P21(serum containing media).

PCT/US01/28861

Figure . Serum anti-p24 Levels at 3 Wks post i.m. immunization of balb'c mice (n=10) with Varying Doses of Several Adgag constructs: (A) MRKAd5gag (through passage 5): (B) MRKAd5 E3\* hCMV-FLgag-bGHpA; (C) MRKAd5 E3\* hCMV-FLgag-SPA; (D) MRKAd5 E3\* mCMV-FLgag-bGHpA; (D) research Lot (293 cell-derived) of Ad5HIV-lgag; and (F) clinical lot (Ad5gagFN0001) of Ad5HIV-lgag. Reported are the geometric mean titers (GMT) for each cohort.



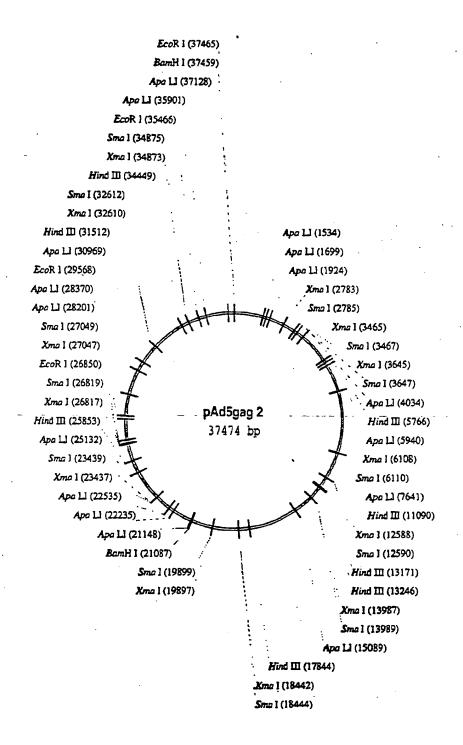


Figure 14

TOCHGRETGG GUIKUUYUU CAGCTGAGGT CTCGACTUCA AGCAGANI 'AA **CCTATAGEA** CATTAGITICA **AATGACGTAT** ATAGTATACC: NGTACATOTA **TCATOTAGAT** COTAC:COT: CCATCCGCAC CTGTGGCCC" ACCACACACAC CGACCTCTCC TCGTCTTGT CCAGATOGT THAN AND THE GACACTCCX C CCATATCAT. GTANTCAAGT TINCTOCATA GGTCTACCAC CITCCCAACCC פנוגנוניניני CICTIONOCO TATCATATO ACAGGTAGG TCTCCACCC AGGCCTTCTG **GGAGGTYCCCT** AGGCTCTGAG TAATGCCCCA ACATCAAGTO GACTUACGGG GATTTCCAAG TCCCGAAGAC CCTCCAGGGA TCCGAGACTC ATTORGOAGG ACCTCCAGGG CTICATGOGG TAACACGTCT TGGAGGTCCC CONTINUENTO TACOPTOTAT ATCCAACATA ATTACGGGGT TOACGTCAAT ACTOCAGTTA TOTAGTTCAC CCTACITIOGC GCATGAACCG CTAAAGGITC **GCAAATTGGGC** COTITACCCG CTCCATAGNA GAGGTATCTT CCCCACCCC TGACOTPITIT ACTGCANANA **GCCATTITICG** COGTANANAGC GCAAATGCAC PROTECTEGNS CANATCACTT GGCAGTCTAG CGGACCTCTG CGGTAGGTGC GACAAACTG OCGANCICATO CATOGONICO COGNITICCC CHICCARIAG TCAGANCINE CATOGOTICOT ATTIGRETICES TANCACACCC OGACCTCTTC ATTIGTECAGA TAGGTAACGT CTGTTTTGAC GTACCCACGA CCCTGCAAAC COGACGITIES CCTGGAGAAG COCCTANCTG ATCCATIGGA GATCANTAAT TATCATTAGT CCCCATTGAC THEOTOACOT AAACACTGCA CTCCCNANG CACCOTITIC GTAAGATTTG CATTCTAAAC GGACTITIGAC CCTGAMCTO ATACTAATCA CCCCCCCCAT GROCCOCOTA ACTROCCAGE CATALANTESIC APPLICACIOGO TIGNACCITICA ACATGACCTT. ATGGGACTTT ATACGCCTICA TGTACTCGAA TACCCTGAAA CTGAGTGCCC ATAATAATAT CCGCCGCCC TANACTIGCCC ATAGGGGTTT OPCATUTAGE FACCOCOCC TATOGCCAAA AACAACTCCG THYTHOAGG GCCATCCACG ACTECTAGATO GCTAAAGCAC CGATTTCGTG CTCCAGCCCT GAGGTCGGGA CCANGGAGGC CONTROL GANCTACCCC 9393333333 טכככאעכטאכ CGGGTTGCTG CUCCOACCIE CGCATTGGCT **NGCIGCCGGGG** TCCCCGCCCC TCATTATTCA CTACTTATTA CCCASCINCANG TOCOCTOCCTA CHICTAACCGA ACCERACGGAT 301 CAGGITETIC TICCAGGIEG TECGALGACG ACCGITETECG TITTAGGITEGG TECACAGGGT TTTTYACAGCA GETTECTANGA AGAAGTACAA TOGGACCOGA CGACCTCTGG AGACTCCCCA CGTCCTTCTA GGACCCCGTC GITTANIAGACA AACTICCACICC AGGINGINGCCA ATICACITATICS ANATISTICGT GCCTGGAGAC CACCATTACTC TYTENTATE CCTGGGGCCAG CITICTIANCITA CACITICCIIGI TACTATTACT AACACATGTA CITAMATTEGG ATATITICAL TATAMCAGA ACTANTANCT CTGGGTTGACC CACCCACTES CITATITIACON TATGCCCAGT TATTTTGGAT TGAMGCGAAT ATGATAATGA TTGTGTACAT CATTTANACC TATTATTA GITTITAGITIG CCCTGAAAGA GITTAGTGAA CCGTCAGATC GAGGGGCCCGG CCCTTGCCAC GTAACCTTGC GCCTAAGAGG TETTETAGTE CGACTECGGA CCACCGTTET CHUTCHCOL GANGATTCAT CCCGACCGTA GCAGGCAGAT CAGTACATCA GGGACTITICC TACAACTGTA CATTGACGTC AATGGGTTGA TTACCCACCT CCCCTCCCAT ACTER YASTIYA ACTIVITIES COSTS TCACACCGCC CCATCTTTCTA CCTACAACAT TRITRITACTE ATARCECTA TATICACIONAL CETTANGECTET CANCCACAAA CATTACCGCC ATGTTGACAT ANTROCCCGC TTACCGGGGG מכשנים אמניים **ו** TCTCAGGGGT TUGGACATGA CACACGTGGT TOUCACACIC TGATGTTGCA CCTTTTACGC CCANANTCCG ACACAATGAG CCGGGTCAM GTAATGGCGG ACTIACGGIA TCAATGCCAT **OTANCTICIONS** CTAMATROCC CATTTACCGG **OCCULTITIOS** CCCANAACC CANANTCANC ATAAAACCTA ACTACAACGT GGCCCAGTTT AGAAGATCAG GCTOGAGACC ACCETATACT ARGETRICARSE GTTTTGGCAC CAAAACCGTG ACCAGAGETE TCAATAATTT ACTTATTAAA MACKEGGAA GCGTTACATA AGGGACTITIC Tecenorana **OTCAATCACG** CCATCGTGAT GOTACCACTA TCATCTCCAA COCMATGTAT CAGTTACTGC TNAMACACACA THYCOGET ACTINCACCITT TAATATATET ATTATATA กรางกราชการ CCCCCTTCAC ATTTRICTOR GTCCAAGANG AAGGCCCAGC GTCTCACCGA GACAAGTGGG GGTCTATATA CTOTTCACCC CACAGTGGCT NACTGANATC TICACTITIAG AGAGTCCACA TTATATTGGC AATATAACCO ATCGAGTTCC TAACGCCAAT ATTOCOGITA CCCTATTGAC GGGATAACTG ATCGCTATTA TAGCGATAAT ATCCCACTT TACCCTCAAA CCAGATATAT CHECOCOGCC ACCCTODCCT CCTTCACTGT TereAddres TACCTCAAGG PTCTTAATTA ACATCATCAA GTAGTAGTGT CATCATCACA GGAAGTGACA AACAATTAAT TETAGTAGTT CCCTOTACAA COGACATGTT TATACATGTA ATCOGGTATA TACCOTCOCA CCGATCCAGC OCCTADOTICG TOGICAGCTG ACCACTCGAC TITOCTOTOA AAACGACACT GANTAAGAGG GTCCACAAAA **TAGCCCATAT** GITCCCATAG CAAGGGTATC COTATTAGIC NTTCACCTCA ATGCCACCCT CCACATOTOT CHTAINTCICC ATATGTACAT CAAGTACOCC GITTCATGCGG GCATAATCAG INCIDCAGT CCCCCACTG CAGGICTETT CACCAGGTOAC DOTOTACACA 1501 100 1101 1201 1301 1401 1601 501 901 601 701 801 301 101 101 201

Figure ISA

## PMRKAITS AND MEREBE

| 1701 | CACCAGGCCA                | TCTCCCCCCG               | GACCCTCAAT                              | CGGACCCACT                    | ACCUPATIONAL TYPE TO THE TABLE THE T | CCTCTTCCCGG               | THETECECTS<br>ANGAGGGGAC                          | ACCACTAGOS<br>TCCACTAGOS               | CATOTICICI                                    | GCCCTGTCTG                               |
|------|---------------------------|--------------------------|---|-------------------------------|--|---------------------------|---|--|---|--|
| 1001 | AGGGTGCCAC                | CCCCCAGGAC               | CTCAACACCA<br>GACTTGHTGT                | TY:C'TGAAC'AÇ<br>AÇGAL;TTGTRG | ACATTA COCASC<br>TCALTA COCASC   | CATCACICTS                | CCATCCACAT  | GCTGAAGGAG                             | ACCATCAATG<br>TGGTAGTTAC                      | ACCITCICAL:                              |
| 1901 | TGAGTGGGAC<br>ACTCACCCTG  | AGGCTGCATC<br>TCCGACGTAG | CTCTTCALGE                              | TradenceAtt<br>AccoractAA     | מכרתבלימיניה<br>התאתקייורניים  | ACATOAGGGA<br>TYTTACTCCCT | CCCCAGGGGCCCCCCCCCCCCCCCCCCCCCCCCCCCCC            | TCTGACATTG<br>AGACTGTAAC               | CTOCCACCAC                                    | CTCCACCCIV                               |
| 2001 | CAGGAGCAGA                | Trocordat                | GACCIANCANC                             | CCCCCATCC                     | CACACCCCT  | ANTCTACAAG<br>TYAGATGTTC  | ACCACCTAGE<br>TCCACCTAGE                          | TCCTC/CCCCT<br>AGGACCCCGGA             | GAACAAGATT                                    | GTGAGGATGT                               |
| 2101 | ACTCCCCCAC<br>TGAGGGGGGTG | CTCCATCCTG               | GACATCANGE                              | AGGGCCCCAA<br>TCCCGGGGTT      | CCTCGGGGAAG  | AGGACTATG<br>TCCCTGATAC   | TGGACAGGTT<br>ACCTGTCCAA                          | CTACAAGACC<br>GATGTTCTGG               | CTGAGGGCTG                                    | AGCAGGCCT** TCGTCCGGA*                   |
| 2201 | CCAGGAGGTG                | AAGAACTCGA               | TCACAGAGAC<br>ACTGTCTCTG                | CCTGCTCGTG GGACGACCAC         | CAGAÁTGCCA<br>GTCTTAGGGT   | ACCCTGACTG<br>TGGGACTGAC  | CAAGACCATC  | CTGAAGGCCC                             | TOGGCCCTOC<br>ACCCGGGACG                      | TOCCALICETO<br>ACOGTOGGA                 |
| 2301 | CTCCTCTACT                | TGACAGCCTG<br>ACTGTCGGAC | CCAGGGGGTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG | GGGGGGCTCTTG                  | CTCAGAAGGC   | CAGGGTGCTG                | GCTGAGGCCA  | TOTCCCAGGT                             | GACCAACTCC<br>CTGGTTGAGG                      | GCCACCATC                                |
| 2401 | TGATGCAGAG<br>ACTACGTCTC  | GGGCAACTTC               | AGGAACCAGA<br>TCCTTGGTCT                | GGAAGACAGT                    | CTTCACGAAG   | AACTRIFICEA<br>TITACACCET | AGGTGGGCCA<br>TCCACCCGGT                          | CATTGCCAAG<br>GTAACGGTTC               | AACTGTAGGG<br>TTGACATCCC                      | CCCCCAAAAA.                              |
| 2501 | GAAGGGCTGC                | TOGAACTICTO              | GCANGGANGG                              | CCACCAGATG                    | AACGACTGCA<br>TTCCTGACGT   | ATCAGAGGCA<br>TACTCTCCGT  | CCGCTTGAAG  | CTOCCCAAAA                             | TCTGGCCCTC<br>AGACCGGGAG                      | CCNCNAGGO:                               |
| 2601 | AGGCCTOOCA<br>TCCGGACCGT  | ACTICCTCCA<br>TGAAGGAGGT | GTCCAGGCCT                              | GAGCCCACAG                    | CCCCTCCCAA   | GGAGTCCTTC<br>CCTCAGGAAG  | AGGTTTCCAG<br>TCCAAACCCC                          | ACCICITICIO                            | CACCCCAGC<br>GTGGGGTCG<br>B                   | CACHAGCAC!<br>GTCTTCGTC!                 |
| 2701 | AGCCCATTGA                | CAAGGAGCTG               | TACCCCCTRIG                             | CCTCCCTGAG                    | GTCCCTGTTT<br>CAGGGACAAA   | GREAAMBACE                | CCTCCTCCCA  | GTAMATAAA<br>CATTITATIT                | GCCCGGGCAG ATCTGCTGTV<br>CGGGCCCGTC TAGACGACA | ATCTOCTOTK:<br>TAGACGACA                 |
| 2801 | CCTTCTAGTT                | COGTECOTAG               | TGTTCTTTGC<br>ACANCANACG                | CCCTCCCCCG                    | TRECTTECTT   | GACCCTGGAA                | GGTGCCACTC<br>CCACGGTGAG                          | CCACTOTCCT<br>GGTGACAGGA               | TTCCTAATAA<br>AAGGATTATT                      | ANTCAGGAAN<br>TTACTCCTTT                 |
| 2901 | TIGCATCGCA                | TTGTCTGAGT<br>AACAGACTCA | AGGTGTCATT<br>TCCACAGTAA<br>Pvol        | CTATTCTOOG<br>GATAAGACCC      | CCCACCCCAC   | CCCGTCCTGT                | GCAAGGTATA  | GCATTOGGAA<br>CCTAACCCTT               | GACMATAGCA                                    | GCCATACGACC                              |
| 3001 | GGATGCGGTG                | GCCTCTATOG               |   | CCGATCGGCG CGCCGTACTG         | AAATGTGTGG<br>TTTTACACACC  | GCCACCCAAT                | AGGGTGGGAA<br>TYCCACCCTT                          | AGAATATATA<br>TCTTATATAT               | AGGTGGGGGT<br>TCCACCCCA<br>Sph1               | CITATGTAGT<br>GAATACATCA                 |
| 3101 | TTTGTATCTG                |                          | 2022000000                              | CCATGAGGAG<br>GGTACTCGTG      | CANCTCGTTT<br>GTTCAGCAAA   | CTACCTTCGT<br>TRACCTTCGT  | THERMACTC AND AND AND AND AND AND AND AND AND AND | ATATTTGACA<br>TATAAACTGT<br>TTGACCTAGG | ACCIOCIATICO<br>TOCCICCATACO<br>AGACCÓTIGAC   | CCCCATISCIC<br>GCCCTACCC13<br>TCCAACCCCG |
| 1026 | OCCCCACGCA                | OTETTACACT               | ACCCGAGGTC                              |                               | פבטעישישישטטט  |                           | GAGATGATGG  | AACTGGATGC                             | TCTGGCACAG                                    | ACCTIGGGG                                |

tique 158

| AACM:TGCAG<br>TTGTCACGTC   | ACCINCIPING IN  | CTACTTCG! 1'  | CTGCACCA  | TCTTCTAGAT!<br>ACAACATCTA  | COMMINITATION COMPRESSION OF THE | GCCCTAAGT     | GCGGGAACAC               | CAGTATCAN                  | GCCTACTTA<br>CCCATCAA''' :                | CCCTCTAGTC  | AAGAGAGCTV:<br>TTCTCTCGAC   | TCGCCGCCCA<br>AGCCGCGGGT  | CCAGGCGGTC<br>GGTCCGCCAT   | GCCACGARCA<br>CGCTRRCCA+ | CGACCGGT   |
|--|---|---|---|--|--|---------------|--------------------------|----------------------------|---|---|---|---|--|--------------------------|--|
|  |   | AAACCTAAAC CT   | TATTITITE ANTIALOTINA<br>ATABARAGO TECTGEACEA                                     | TOCOGGING TO ACCOCCCOCC AC   |  | GTATAGGGAG GC |                          |                            |   |   |   |   | CCAAGCAGTT<br>GGTTCGTCAA<br>CCGCAGTAGT                             | CCCGCCTCC                | CCCCGACGC  |
| CTFFCCTSAG   | ACTTANTGTC  | CCAGACTCTG  | CCCAGGACAC  | AOCTTCATGC   | CCGTCCGGGA<br>TGTTCCCAGC   | ACAAGGGTCG    | ACCCACCITIC              | 2 ATATTTCTGG<br>TATAAAGACC | ACCCOTATA TOCTACOTAG ACCCCATATT ACCAGGTAG | PETACETEGO GOCCONTGAA GAAAACBIT TECUASUING<br>AGATGGACGC CCCCCTACIT CTTTTGCCAA AGGCCCCATC | CTOGGNAGNA NOCAGOTTCC TGACCAGCTG CGACTTARCG CARCAGTTAG GCCCGTANAT CACACCTATA ACCGGGTGCA ACTOGTAGT GACCATAGA GGCCATAGA TGGCCGACGT TGACCATCAA GACCCTTCTT TCGTCCAAGG ACTCGTCGAC GCTGAATTGC CTCGGCCACC CGGCATTTA GTGTGGATAA TGGCCGACGT TGACCATCAA | GANTECETATO TITTECCTES CEMANICOGO CAGANGOGO CITCANGETACOCO GIETTAGOCO GIETTECOCO SINI                           | T GAGCOTITICA<br>A CTCGCAAACT<br>C TITICGCIGTA                     |                          |  |
| ACTÓNCTTTG<br>TGACTGAMAC   | TEACCECTORA<br>ACTEGRICECT                                | AAATAAAAA<br>TTTATTTTTT   | CGGTCGTTGA<br>GCCAGCAACT<br>PSI   | ACCACTVICAG<br>TGGTGACGTC  | GATTRICCAGG<br>CTAACGGTCC<br>AGGTTRGGCTA   | TCCAACCGAT    |                          |                            | TCCCCTATAA                                | CCCCCTACT   | CACACCTAT   | TTTTCCCTG<br>AAAAAGGAC<br>Spih  |  |                          | NGTCCCACT  |
| היהמתאדועיות<br>הייבריויאת:אכ  | TTKIGATTCTT<br>AACCTAAGAA                                 | TTTAAAACAT<br>AAATITIKITA   | CCAGCCTTCT  | TYSAGGTAGC<br>ACCTCCATCG   | CATCGTTCGA   | CACATAAAA     | TCATGTAGCT<br>AGTACATCGA | CCCGCCGCCG                 |   |   | GCCCSTNANT  | GACTCGCATG  | TITIONGACCG TCCCCCTTAG AAACTCTTAIC AGGCACATC                       | ATATAGAGAN; CAAAGAGGG    | CCTCATCACC GTACACTURAS GGAGGCACTURAS CATCAGACTURAS ATCAGACTURA CATCAGACTURA CATCAGACTURA CATCAGACTURA CATCAGACTURA CATCAGACTURA CATCAGACTURA C |
| איונימאמאא, ו  | TITIVALCACIAN TITVIGATICITY ANACCIANGAN                   | CCCAATGGGG TITTAAAAGAT<br>GGGITTAGGGC AAATITFIGTA   | AGGCCCCT GCCCCAGA GCCAGCACT TCCGCGCCT GCCCCAGA GCCAGCACT PCGCCAGCCAGCT FILE Pall  | GACACATORGICA CONTRACTORICA CO | ATCHETTICA GTAGGAAGET<br>TACMGAAAGT CATCGTTCGA<br>GCANTINGGA CHETATTITT  | CGTNTAACCT    | GGGAAATTTG               | ATTACK CCAC<br>TACK CGGGTG | GCCACAGCAAG                               | CCCCTAGTAC  | CAGECGRITIG   | GCATGTCCCT  |  | TATAGAGGAG               | CCTYTHYAGG   |
| PSH  CHARLES CCARTEGORG CONGANICACI ACTONOCITICS  CHARLES CONFORMAN CONGONIANO TONOCITICS  CHARLES CONFORMAN CONGONIANO TONOCITICANOCITICS  CHARLES CONFORMAN CONGONIANO TONOCITICANOC | TENCEGRANCE TITTERCACEA L'ESTANGAL TENCECCERTE ACTEGRECET | THECTREGET CECANTIFIC TITANAMENT ANATANAMAA<br>ANGGARATIA GRUTTAGEC AMPLIFICIA PETATITITI | CHARACTER AGGECTION CENTRAL CONTROLLES GEGENAL GEGENACT GETCHENDA GECNGENACT PAIL | AGATACATGG GCATAAGGCC GTCTCTKGAGG TKRAGGTAGC TCTATGTACC CGTATTCGGG CAGAAGACC ACCTCCATCG  |  |               | CCCTCCACTT               | AATGATGGGA<br>TTACTACCGT   | TTTACAAAGC<br>AAATGTTTTCG                 | GITCAGATOS OXYANTCATO CAAGTCTAGC CCCCTAGTAC   | ACCAGOTTCC TONGCACCTG CGACTTACCG TCGTCCAAGG ACTCGTCGAC GCTCAATGCC   | PAN  COCCOCCO CATCCCTCAG CAGGAGGCC ACTTCGTTAA GCATGTCCCT  GCCCCCCCG TCAAGGACTC GTCCCCCCGG TCAAGGAATT CGTACAGAAA |  |                          |  |
| CGCCGCTTCA (   |   |   |   | AGATACATGG   | CGACCTGCAC   | CATACGTCAGG   | ACAGTGTATC<br>TGTCACATAG | ATTEGREEAT                 | ATACCICATT<br>TATCCICATA                  | CACCCTTTGA  | TGACCAGCTG<br>ACTCGTCGAC  | CATCCCTGAG CAGGAAAGCC<br>GTAGGGACTC GTCCCCCCGG  | OCCUPAGORIO TECTECCIAGO GINOCANOT CACINECITE ANGIACOTEC CITECTITES | CTACGGCATC               | CAGGGTCATG TCTTTCCACG<br>GTCCCAGTAC AGAMAGGTCC   |
|  |   |   |   | CTOGATOTIC   |  | CCCTACCCAC    | AACCACCAGC               | THINCCATCC<br>AAAAGGTACG   | TCACATCGTC                                |   | AGCAGGITICC<br>TCGTCCAAGG   | CATCCCTGAG<br>GTAGGGACTC  | TTCTTGCAAG<br>AAGAACGTTC   |                          | CCAGACGOGC CAGGGTCATO  |
| Pali<br>mmmme<br>TTOOMGACTO CAGCTCCGC  | CTTCCCOTTC /  |   |   | AAAGGTGACT   |  | CCCAATICGA    | TOTTOTOCAG<br>ACACACATO  | ACCTCCAGA                  | TOTTCCAGGA                                | CCTCACAGAT  | CTOOGNAGAA  | Psil<br>CACCTCCCT<br>GTCCACCOCA   | OCCUPATOCAG<br>CGCTATCOTC  | CCACAGCTCG               | CCAGACGGGC   |
| 301  | 101   | 1501  | 1601  | 3701   | 3801   | 3901          | 4001                     | 4101                       | 4201                                      | 4301  | 4401  | 4501  | 4601   | 4701                     | 4801   |

Figure 15c

# рмиклабара мек682.

| 4901        | CCACGCGAAC   | AGGCTOSTCC     | TCCTCCTCCT<br>ACCACCACGA  | CAMPACATINE C  | CUCTUTURE CO         | CCTOCCCCTC CONTROL | CCCCTCCATC    | CATTICACCA<br>GINAACTOOT | TCCTCTCATA<br>ACCACAGTAT | GTCCAGFCFC<br>CAGGTCGGGF |  |
|-------------|--|----------------|---|----------------|----------------------|--------------------|---------------|--------------------------|--------------------------|--------------------------|--|
| 5           | , ŧ  |                |   |                |                      |                    |               | TOAGGCGTA                | GAGCTTAGGC               | GCGAGAAATA               |  |
| Tone        | Production A   |                | המנים המארות היינות המינות br>המונות המינות |                |                      |                    |               |                          | CTCGAACCCG               | CCCTCTTTAT               |  |
|             |  |                | 200200000   |                |                      |                    |               |                          | TCGGGGTCAA               | AVACTCAGGITT             |  |
| 21015       | GCTAAGGCC  |                | אנטכנינינינין   |                |                      |                    |               |                          | AGCCCCAGTT               | Tringencea".             |  |
|             |  |                |   |                | _                    |                    |               |                          | Usit 1071                |                          |  |
| 5201        | Tecestage  | TITTICATOC     | GITTICITACC   | TCTGGTTTCC     | ATMACCOCT GTCCACGCTC |                    |               |                          | TOTOCOCOTA               | TACAC:ACTTV              |  |
| ı           | В  | AANAACTACG     | CAMBANTOG   | AGACCAAAGG     | TACTCGGGCA           | CACGTCCGAG         | CCACTGCTTT    | TCCGACAGGC               | ACAGGGGCAT               | ATCTCTGAM                |  |
|             |  | Xhol           |   |                |                      |                    |               |                          |                          | •                        |  |
| 5,101       | AGAGGCCTGT   | CCTCGAGCGG     | TOTTCCGCGG  | TECTECTEGE     | ATAGAMACTC           | GOACCACTCF         | CACACAAAGG    | CTCGCGTCCA               | <b>GCCCAGCACO</b>        | AACKONCICTA              |  |
|             | TCTCCGGACA   | OGAGCTCGCC     | ACANGGICTICC  | ACCACCACCA     | TATCTTTGAG           | CCTGGTGAGA         | CICIGITICC    | GAGCGCAGGT               | ccconconce               | Treeteegat               |  |
| 5.401       | Actional   | GTAGCGGTCG     | TTGTCCACTA  | GGGGGTCCAC     | Trachecage           | <b>OTICTCAAGAC</b> | ACATGTCGCC    | CTCTTCGGCA               | TCANGGAAGO               | TOAITTOSTIT              |  |
|             | TCACCCTCCC   | CATCGCCAGC     | AACAGGTGAT  | CCCCCAGGTG     | AGCGAGGTCC           | CACACTTICTE        | TOTACAGOOG    | GAGAAGCCGT               | AGTICCITICC              | ACTANCCANA               |  |
| 5501        | CTACATATAG   | GCCACGTGAC     | COGGRETICC  | TGAAGGGGGG     | CTATAAAACG           | GCCTCGGGGC         | acentearce    | TCACTCTCTT               | CCGCATCGCT               | GTCTGCGAGG               |  |
| ,           | CATCCACATC   | COGTOCACTO     | GCCCACAAGG  | ACTTCCCCCC     | GATATITICS           | CCCACCCCCG         | CCCAACCAGG    | ACTGAGAGAA               | GGCGTAGCGA               | CAGACGCTCC               |  |
| 5601        | CHURCHCHE  | CXXXXTCAGTA    | CTCCCTCTGA  | AAAGCOGGCA     | TCACTTCTGC           | GCTAAGATTG         | TCAGTTTCCA    | AAAACGAGGA               | COATITIONTA              | THE ACCENCE!             |  |
|             | CONTEGACAA   | CCCCACTCAT     | GAGGGAGACT  | THEGGEGGT      |                      | CCATTCTMC          | AGTCANAGGT    | TITITICINCT              | CCTAMACTAT               | ANGTGGACC'               |  |
|             |  |                |   |                |                      |                    | HIMBI         |                          |                          |                          |  |
| 1073        | Construction & Constr | Cartifold Acad | tenasociant.  | CCATCTCGTC     | ACMANAGACA           | ATCTITITION        | TOTCAAGCTT    | TOTCAACCTT GGTGGCAAAC    | GACCCGTAGA               | CRATCOTTICA              |  |
| 1015        |  | CCATABACTCC    | CACCGGGGTA  | GCTAGACCAG     | retrireser           |                    | ACAGITCGAA    | CCACCGITITIG             | CTOOCCATCT               | CCCCCAACCT               |  |
|             |  | }              |   |                | Pyul                 |                    |               |                          |                          |                          |  |
| 100         |  |                | Car Braziliania   | Calendalatical | CGATCGCGC            | GCTCCTTGGC         | CGCGATGTTT    | ACCIDEACOT ATTERCEGEGE   | ATTCGCGCGC               | MCGCACCCK.               |  |
| 1000        | GICOTTGAAC   |                | COTCCCAAAC  | CAMANCAGO      | CCTACCGCC            | CCAGGAACCG         | GCCCTACAAA    | TCGACGTGCA               | TAAGCGCGCG               | TICCOTCCCJ               |  |
| 5001        |  |                | OCOCTCGTCG  | GGCACCAGGT     | פכעכניכניככע         | ACCICCOCTTO        | TGCAGGGTGA    | CANCOTCAAC               | OCTOOTGOCT               | ACCTOTOCG                |  |
|             | GTANGCCCTT   |                | CCCCAGCAGC  | CCGTGGTCCA     | Cardidenca           | TCACGCCAAC         | ACGICCCACT    | GFFCCAGFFG               | COACCACCOA               | TOOMGAGG!                |  |
| 6001        | GTACACACAC   | GFTGGTCCAG     | CAGAGGCGGC  | COCCCIPICACE   | CGACCAGAAT           | GENERALAGE         | GGTCTAGCTG    | CONCINCATOR              | occoconcia               | CGTCCACGGT               |  |
| ·<br>><br>> | CATCCGCGAG   |                | GTCTCCGCCG  | GCGGGMCTC      | CCTCOTCTTA           | CCCCCATCCC         | CCNGATCGAC    | GCAGAGCAGG               | CCCCCAGAC                | OCAGGTACCA               |  |
| 6101        | AAAGACCCCC   | -              | OCCCCTCGAA  | GTAGTCTATC     | THECATECTT           | GCAACTCTAG         | ממכנותכנופכ   | CATGCGCGGG               | COCCANGCCC               | GCCC**CCTA*              |  |
| 4           | T-T-C-10000C   |                | CGCGCAGCTT  | CATCAGATAG     | AACCTAGGAA           | COTTCAGATC         | GCCAGACGACG   | GTACCCCCCC               | accericace               | CCCCACCATA               |  |
| 6201        | CACIFICAGIO  |                | TOCCATGREG  | TCGCTTAGCG     | CGGACGTA             | CATGCCCCAA         | ATGTCGTAAA    | CCTAGAGGGG               | CTCTCTGAGT               | ATTCCAAGAT               |  |
|             | CCCAACTCAC   |                | ACCGTACCCC  | ACCCACTICGE    | GCCTCCGCAT           | GTACGGCGTT         | TACAGCATTT    | GCATCTCCCC               | GAGAGACTCA               | TAAGGTTCTA               |  |
| 6101        | ATCTACCCTA   | GCATCTTCCA     | CCGCGGATTGC   | TOCCOCOCAC     | GTANTCRIAT           | AGTTCGTGCG         | ACCORDECTAG   | CACCINCCCGA              | CCGAGGITTGC              | TACGGGGGG                |  |
| ·<br>·      |  |                | GGCGCCTACG  | ACCGCCCCTG     | CATTAGGATA           | TCAAGCACGC         | Tecenedene    | CTCCAGCCCT               | GCCTCCAACG               | ATGCCCGCC                |  |
| 6401        | CTGCTCTGCT   | COGRAGACTA     | reference   | GATGCCATGT     | GAGTTAGATO           | ATATGGTTGG         | ACCCTCCTCCAAG | ACCITICAAGE              | reaceretes               | DAGACCTACC:              |  |
| ·<br>•      | GACGAGACGA   |                | AGACGGACTT  | CTACCCTACA     | CTCAACCTAC           | TATACCAACC         | TOCOACCTTC    | TOCAACITICG              | ACCOCAGACA               | CTCTGGATG:               |  |

Figure 150

# PMRKAdSgrag MIR682

| GTCCAGRGTT TCCTTGATGA<br>CAGGTCCCAA AGGAACTACT<br>ATCCGAAGC CGTCGACTT<br>TAACCTTTGG GCARCCGAR |  |                          |                          |                          |                          |                          |  |                                 |   |   |   | GAGCTOTCCA TOGTCTGGAG<br>CTCGACAGOT ACCAGACCTC | GATACCTAAT TITCCARGGIC.                                | SOCIONE TECHTRICATE   |
|---|--|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--|---------------------------------|---|---|---|--|--|---|
| CCCCCCTCATA<br>CCCCCCTCAT<br>GTACTCTTOG   | ACCCCCTATO<br>TCGCCCATAC   | COCATCCGCC               | AGCCATAAAG<br>TCCGTATTTC | TOCCCCACAA               | CCACACATIC               | AAACTGGCGA               | GCGGCAGTCA   | CCTARGTGAC<br>GCATCCACTG        | GTAGAAGTCC<br>CATCTTCAGG                    | ACCTGACGAC<br>TGGACTGCTG                      | CHARCTECTIC<br>GACCGACGAG                                       | CCCCAGATGG<br>CCCGTCTACC                       | AGATECAGGT<br>TETAGGTECA                               | GACTACEAT CHECCENCO GOOGGAGG COCOCOCAC CTGATGCTA CACCCCCAC CCCCCACCC CCCCCCCCACCCACCC |
| Trichingteth<br>Aceteracht<br>Getetetren  | THETACHOUS   | TCAGTGTCGT<br>AGTCACAGCA | TTY COGOGGG              | CTTCATCTTC<br>CAACTACAAC | CTCAGCCCGT               | GAMMAGRICCE              | TAGGTCTCGC   | GAGAGATGTA                      |   | CACGAGGITIG                                   | CCTTOACCGT<br>GRANCTURICA                                       | TCACAACATC                                     | 39555579657<br>55555555579<br>11                       | יייייי<br>ייז יהניהכיאיכל<br>משכמיניםכינים  |
|   | ACKTATCCCTT<br>TCCTTACGGAAA  | GTATTTGAAG<br>CATAAAGTTC | AAGAGTATCT<br>TTCTCAFAGA | CCTCAAAACCC              | TTCACGOCAG<br>ANGTECCETE | ACGTCGTCGC<br>TCCACCAGCG | GGTTCGCGGC   | CCAAGTATAG<br>GGTTCATATC        | TOCCTATTGA<br>ACCGATAACT                    | GTACATCCTG                                    | GGCTGGTTGT CCTTGACGGT<br>CCGACGAACA GGAACTGGGA                  | CCGCCCAACT TCACAACATC                          | GACCKIGTCAG GACCCCFGCT<br>CTGCCCAGTC CCGCGCCGA<br>Kprl | GACTACGGTA<br>CTGATGCCAT  |
| ACTOSTICIONS O  |  | TGAGGTACTES              | CTGTARCCATC              | AGCACGATCT               | AGGTGAGCTC<br>TCCACTCGAG | TAGCATTTGC<br>ATCGTAAACG | TCCCATCCAA<br>AGGGTAGGTT   | AGGCCCCAT                       | ATTOCAVACAG<br>TAACCTCCTC                   | TGCACCACCT<br>ACGTGCCCCIA                     | CTTCTACTTC  | GCGCGCGCGT                                     | ACCTCGCATA<br>TCGAGCGTAT                               | מסטטטטטטט<br>מסטטטטטטטט   |
| CGCAGCTTGT TCACCAGTTC GGCGGTGACGGGGGTGCGGACGACGGGGGGGGGG                                      | GTTGACGGCC 1   | ACCATGACTT TAGTACTA      | CCCGCTTCCA               | באמשבפטבט פאבבבפב        | ACTICCTCGT TCAAGGAGGA    | CACAGGGCAT               | TTCCCAGCGG   | TOCTTCCCAA                      |   |   |   | AGATGTCCGC<br>TCTACAGGCG                       | CTCCACGTTT   | AGCCGCATC<br>TCCGCGTAG  |
|   | TGTAGAACTG C   | CCACAGGGAC .             | GGATTTGGCA (CCTAAACCGT   | TOTTANTTAC ACMITTANTO    | CAATTITIA                | CTCCACAGGT               | OCCIONATE COCCIONATE C | CCCGTGCTCG                      | AACTGGATCT CCCGCCACCA TYGACCTAGA GCGCGCTGGT | OTOCOCACTA CTOCOACOG<br>CACGCOTCAT GACCOTCGCC | TOGCGGGTTT  | CCCAAAGTCC                                     | GCGGGAGCTC<br>CGCCCTCGAG                               | GOCTTICANG  |
|   | GAGCCTAGCA CONCRETE TO CHECK AND CHECKS AND CONTROLL TO CHECKS AND C | TGACCCCAAA O             | PHYSGAACGC AAACCTTGCG    | TCGGAACGGT<br>ARCCTTGCCA | TCATCCTTCC ACTACCTTCC    | CHCCANTCAC               | TAGAAGGTAA   | CCAGCATGAA<br>GGTCGTACTT<br>Pvd |   | TTOTAAAAC<br>AACATTTTTG                       | OCCCCTCGCC TUGCGGGTTT GGCTGGTGGT COORDAGGG ACGCCCCAAA CCGACCACA | GCCCCCCAG                                      | GREAGGTCAG   | TREFFICENCE COCCUTCGAT<br>ACCAACCACC GCCCCAGCTA                                       |
|   | CTAACOCTAA C   | CTCCACACCC               | CCOTCCCCTT               | TCCCGGCACC<br>AGGGCCGTGG | CCCTACGOGA               | GOTTOGAAGC               | CCACTACOTC   | AACTTCATGA                      | CHACGETEGG CHACCOGRAG                       | GTGCTGGCTT                                    | CCCTTAAACT  | OGACCACCAC<br>CCTOGTOGTO                       | CTCCCGCGGC   | TESTIGGIGG  |
| 6501  | 6701   | 6801                     | 6901                     | 7001                     | 7101                     | 7201                     | 7301   | 7401                            | 7501  | 7601  | 1011  | 7801   | 7901   | 8001  |

## PMRKAd5gag MBR682

| 101   | ATCCATCTAA               | < t           | Acceptate geometric  | CCCCCCANCO ACCENTANCO  |                               | CONTRACCON CASSARANAS GGCAGGGGA CONCARIOCO GONGGGGGGN<br>ANCONSAGA ANCONONICO CONCOCOGA GCAGGGGGGG CAGGGGGGN | COCCAGACCO (           | GCCAGGGGA CCGTCCGT                          | CONCORRECT         | כמכמנומכבנה:                           |  |
|-------|--------------------------|---------------|----------------------|--|-------------------------------|--|------------------------|---|--------------------|--|--|
| 100   | ACCAPTION I              | - 0           |                      | GUGAAUGUA  | LUNGARA                       |  |                        | GCCTCTGCGT                                  | GAAGACGACG         | מוכבנגאנוגיא                           |  |
| 4     | TCCTCGACCA               | ט פ           | ATCCAACCIAC          | CGCTTCCGCT   |                               |  | ACTITAGACCG            | CGGNGACGCA                                  | crictectec         | CCGCCCACI                              |  |
| 101   | GCTTGAACCT               | G             | TEGNESAGART          | CANTITICGGT  | CHECHTGAGG                    | מנינינונינינוניניני  |                        |   | CCTGAGTTGT         | CTTKATAGG:                             |  |
|       | CGAACTTOGA               | CITICICICA    | AGCTGTCTTA           | GTTAAAGCCA   | GTTAAAGUEA CAGEAACTIGE        | CUUCUCUS   | CCITITITAGAG           | GACGIGCAGA                                  | GGACTCAACA         | GAACTATCC:                             |  |
|       |                          |               |                      | Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Ma<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Mariana<br>Ma<br>Ma<br>Ma<br>Mariana<br>Ma<br>Ma<br>Ma<br>Ma<br>Ma<br>Ma<br>Ma<br>Ma<br>Ma<br>Ma<br>Ma<br>Ma<br>Ma |                               |  |                        |   |                    | · · · · · · · · · · · · · · · · · · ·  |  |
| 401   | GATCTCGGCC               | ATGAACTOCT    | CONTCICTING          | CTCCTVGAGA   | כזיביכיויהמאפא זכיויכיניקיהיי |  |                        | GCGAGGTCGT                                  | TOGRAMITACO        | GCA.CATON:                             |  |
|       | CTAGAGCCGG               | TACTTGACGA    | GCTAGAGAAG           | CACHICT  | ACACKOCGCAG                   | GCCCAGCGAG   |                        |   | ACCITITACE         | ברנאיז אר זיר.                         |  |
| 501   | TOCCHOAGO                | COTTORORCC    | TCCCTCGTTC           | CAGACGCGGC   | TUTAGACAC                     | פנ.ככככבונונפ  |                        |   | CACCTOCOCO         | ACATTICACE "                           |  |
|       | ACOUNTINC                | ٠             | AGGGAGCAAG           | GTCTGCGCCA   | ACAMETRICATE                  | CHAGGGGANAGC   | CCTAGCGCCC             | GCGCGTACTO                                  | OTGGACGCGC         | TCTAACTCG.                             |  |
| 109   | CCACGTGCCG               | GCCGNAGACG    | <b>GCGTAGTTTC</b>    | GCACGCGCTG   | AAAGAGGTAG                    | THEACCEITES  |                        |   | AAGAAGTACA         | TANCCARC:                              |  |
| !     | GOTGCACGGC               | _             | COCTICIOS COCATCAAAG | CGTCCGCGAC   | TTTCTCCATC                    | AACTCCCACC   | ACCECCACAC             | AAGACGGTGC                                  | TTCTTCATGT         | ATTGGGTC(X)                            |  |
|       |                          | ī i           | ECON V               |  |                               |  |                        |   |                    |  |  |
| 701   | TCCCMCGTG                | G             | INTEGRICA TATCCCCAN  |  |                               | CCTCCTAGAA   |                        | ANGITICARAN                                 | ACTORGAGE          | GCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC |  |
|       | AGCOTTGCAC               | : CTANGCAACT  | ATAGGGGGTT           | CCCCAGITICC  | GCGAGGTACC                    | GENGCATCIT   |                        |   | -                  | 4.1                                    |  |
| 108   | ACCURITABLE              | r cerectecas  | AAGACGGATG           | AGCTCGGCGA   | CAGTGTCGCG                    | CACCTCCCGC   |                        |   | renterici          | TCAATCTCE:                             |  |
| •     | TOCCAATTICA              |               |                      |  |                               | GTGGAGCGCG   | AGTITICCGAT            | GTCCCCGGAG                                  | AAGAAGAAGA         | ACTTAGAGGA                             |  |
|       |                          |               |                      |  |                               |  |                        |   | SAN CANADA         | 1                                      |  |
| 100   | C. Parter B. T. B. B. B. | المعاسيديدي   | للحاماد للمادعان     | CTOCCOCCOG   | TGGGGGGAGGG                   | GGGACACGGC   | GGCGACGACG             | GGCGACGACG GCGCACCGGG AGGCGGTCGA CAAAGCGCTC | AGOCOCOTICGA       | CAMBEGETE                              |  |
| 7 0 0 | CHICCAIRM                | -             |                      |  | אנכניננונכני                  | CCCTGTGCCG   | CCGCTGCTGC             | COCOTOGCCC                                  | TCCGCCAGCT         | TCCGCCAGCT, GTTTCGCGAG                 |  |
| 150   | CARTERITUE               | _             |                      |  | Graciance                     | Tricacadaa   | GCCCAGTTOG             | ANGACGCCOC CCGTCATGTC                       | CCGTCATGTC         | CCCGGTTATCG                            |  |
| 7     | CTAGTAGAGG               |               |                      |  | כפנפנכנאטט                    | AGAGCGCCC  | CGCCTCAACC             | TTCTGCGGCG GCCNGTACAG                       | GCCAGTACAG         | GCCCANTACC                             |  |
|       |                          | •             |                      |  | CGATCCATCT                    | CAACAATTGT TGTCTAGGTA  | TUTCHAGGTA             | CTCCGCCGCC GAGGGACCTG                       | GAGGGACCTO         | AGCGARTCIT                             |  |
| 101   |                          |               |                      |  |                               |  | ACACATCCAT             | DAGCCGCGG CICCCTGGAC                        | CICCCIOGAC         | TCGCTCAGG                              |  |
|       |                          | •             |                      |  |                               |  |                        |   |                    |  |  |
| 100   |                          | 2 ATTOTABABLE | CTCTCGACAA           | ACCCCTCTAA   | CCAGTCACAG                    | TYCCAARGTA GGCTGAGCAC CGTGGCGGGC GGCAGCGGGC  | GGCTGAGCAC             | CGTGGCGGGC                                  | ООСУФСОООС         | OCCCCTCCCX:                            |  |
| 707   | GTAGCTGGCC               | • •           |                      |  |                               |  | CCCACTCGTG             | GCACCGCCCG                                  | CCGICCCCC          | CCGCCAGCCC                             |  |
|       |                          |               |                      |  |                               |  | Manuscon               |   | -                  |  |  |
| 101   | CHARTITA                 | 3 GCGGAGGTGC  | TGCTGATGAT           | GTANTTANAG   | TAGRECIANT                    | TYAGACGCC  | GATOCTICAC             |   | יוסיוכביזינאני     | וררואורווויר                           |  |
|       | CAACAAAGAC               | _             |                      |  | NTEXTRICAGA                   | ACTCTGCCGC   | CTACCAGCTG             | remediast                                   | ACAGGAACCC         | AGGCCGGACG                             |  |
| 1070  | TALL DURING A CALL       | _             |                      | GCTTCGTTT  | GACATICAGG                    | CARCTETTO  | TAGTAGTETT             | GCATGAGCCT                                  | TICTACCOGC         | ACTIVETICE.                            |  |
|       | ACTIVACION               |               |                      |  |                               | GIECCAGAAAC  | ATCATCAGAA             | CCTACTCGGA                                  | AAGATGGCCO         | TGANGNAGA                              |  |
| 1020  |                          |               |                      |  | ניבניביניניני                 | CACTTINGCIC  | CTACATIGACA            | CCCTCTTCCT                                  | CCCATGCGTG         | TGACCCCGAA                             |  |
| 100   | CICCIICCIC               |               |                      |  |                               | CTCAAACUGG   | CATCCACCGC             | CCCACNAGGA                                  | CCGTACGCAC         | ACTGGGGCT I                            |  |
|       |                          |               |                      |  | CCTCGGCTA                     |  | ATATORICHO CTRICACCTOC | GTGAGGGTAG                                  | <b>ACTOR/ANGTO</b> | ATCCATGFCC.                            |  |
| 1006  | SCCC CARG                |               |                      |  |                               | TATACCGGAC   | GACGTGGACG             | CACTCCCATC                                  | TCACCTTCAG         | TAGGTACAGG                             |  |
| ٠.    |                          | _             |                      |  |                               |  |                        |   |                    |  |  |

# PMRKAd5qag MER682

| PACC<br>ATGG  | AGAC:<br>TCTC   | 1000:<br>1000:<br>1000:   | JACCT!   | AGIY:                              | TICA('                   | AVAC: :  | NCGC<br>TGACG | ACTIVIC<br>RETRUC | AGAK's:<br>ACGAG         | TCCTC<br>AACCG           | TIGGC                    | GCTTTANCAN<br>CGAAATTGTI         | ATCGTTCGGC            | GCGACCGACG  |
|---|---|---|--|------------------------------------|--------------------------|--|---------------|-------------------|--------------------------|--------------------------|--------------------------|----------------------------------|-----------------------|---|
| TCGGTGTATC<br>AGCCACATGG  | CCCCCATCTC  | GCCGGCGGC . AGGCGCGCGC . TCCGCGCGC .  | CGGACGACCT   | CTCACGAGGA GCATTAAGIY:             | CGTAATTCAC               | CCCCCANG   | COTCTACGC     | CACCANCIVIC       |                          | CCCCCTCCTC               | TG CGCGC                 | CCTTTANCAN<br>CGAAATTGTT         |                       |   |
|   |   | TCCAGGIGAT AGGICCACTA CIGGCCGGIC GACCGGCCAG   | GOTATCATOR   | CHCTTGCCCC                         |                          |  |               |                   |                          |                          | ACTCOGGC                 | AAAGTTT                          |                       |   |
| ANCHERACTAG TTAACGGTCT GGTTACCCGG CTGCGAGGGGTTTGGTCTCGGGGCTCTCGGGGGGGGGG                  | ACCAPATANCT GATATCCCAC CAAAAAGTGC GGCGGCGGCT TXXTCCCATAA CCATAGGGTG GTTTTTCACG CCGCCGCGG  | TACCTOGACA ATGGACGCT TCGGACGCT AGCCCTGCGA   |  | TOCCACCTON  ACCOTOCACT  COSTAGOSTO |                          |  |               |                   |                          |                          |                          | CCCTCANCCA COAGATTAAC CCCCTAATTG |                       |   |
| TTAACGGTCT<br>AATTGCCAGA  | GGTATCCCAC<br>CCATAGOGTG  | CCGCTACTATA TCCGTAGATG CCGCTACTATA AGGCATCTAC CCCCCAAAAAG TGCTCCATTG CCCCCAAAAAG TGCTCCATTG CCCCCAAAAAAAAAAAAAAAAAAAAAAAA |  | ANCECAGGING<br>THERESTECAG         |                          | CTCAGAGCCT   |               |                   |                          |                          |                          |                                  | T GAAACATTCG          |   |
| AACGGACLAG<br>TTGCCTOGTC  | ACCAGGITAGT C<br>TXXTCCATCA C   | CCCCTACTAT CCCCTACTAT CCCCCAAAAAG   |  |                                    |                          |  | TCGGGGANA     |                   |                          |                          | AGCGCTCGC                |                                  | TAGACACCT  TAGACACCCT |   |
|   | GCTTCAGGGGG   | CCGCTCTAGA AGGITGIANT CCGCTCTAGA AGGITGIANT CCGTTCCTAGA TGTTGGCAG   | ACCOSTRACT TEGECECOTES   | GGCANTACKGC<br>GCCANTACHG          |                          |  |               |                   |                          |                          | ATGCCCTCAA               |                                  |                       | TOTTRETERO  |
| GATSTANGTTGC  | CCTACTECTT CCTACTECTIC<br>GCATCACTA CCTTECACTOC<br>Null   | CCGCTCTAGA AGGTTGTATT CCGCTCTAGA AGGTTGTATT CCGTTCCTAGA   | GCCAAGGTCT ACAMMACAMATA<br>ACAGGCTYTA AGGGTGTGACT<br>TCTCGGACAT TGGCCGGTAA | GTCTAGGTAC                         | GCGCTAGCTT               | CCAAGOGTTG   |               |                   | -                        |                          | GAGCTGCCAAC              |                                  |                       | CACAGCAGAGA   |
| HATIGGGG CGTGTTGATG   |   | GOCTCCASE<br>CCCGAGGCCC<br>GTCGCGGACG   | CAGCGCCTISC GCCAAGGTCT  GTGCAAAAGG AGAGGCCTTTA  CAGGTTTTCC TCTCGGACAT      | COCCACCACC                         | CCCCCACCA                | GCCANTAAAA   |               | AGAGCAMGAG        |                          |                          | CCCANGGCGC               | CACGTOGING                       | CECCICGAGGA           | ACCRETICCT TATACTICAG CACAGEAGES TCGACAAGGA ATATCACGTC GTCTCCTCCC |
| GGTATGCCC COTGTTGATG GTGTAAGTGC AGTTGGCCAT<br>CCATAGGCG GCACAACTAC CATATTGAGG TCAAGCGGITA | Shell | AGGOTOUCCG<br>TCCCACCGGC<br>CGCGCGGAAA  | ACGCCCCTTP  Xbal  COTCTAGACC  CGAGATCTGG                                   |                                    | Trechooede<br>Anosteeded | TOTAGCCGGA<br>ACATCGCCCT   | TOCANGACCC    | AGCAGCGGCA        | TOGTUATTAC<br>ACCACTAATG | CCANGGOTGC               | GAAAGTICCA               | COCOCOCOCOCA                     |                       |   |
| ACAAGCGGF G   | TCAGACGCGA C  |   | CACCACCTCC  AATCOTIGAC  TTAGCAACTG   |                                    | TTTGGCTTCC<br>ANACCGNAGG | CONOCONOCO CON CONTRACTOR CON CONTRACTOR CON | CTCCCCGTCA    | CCCCCTCCTC        | COCCACCACA               | TRABCOCCAC<br>ACTCOCCCTO | ATGCGGGATC<br>TACGCCCTAG | GGATTAGTCC<br>CCTAATCAGG         | CCACCITOCOF           | CTCATGGCGC<br>GAGTACCGCG  |
| 9701  | 9801  | 10001   | 10101  | 10201                              | 10301                    | 10401  | 10501         | 10601             | 10701                    | 10801                    | 10901                    | 11001                            | 11101                 | 11201   |

| 11301 | TECATITICAL A | AMCATECTG                | CAGAGEATER  | TRICTICACION          | נוננו. איצידונו        | AGENTAGGETO                  | ACAMPGTOGF                        | CGCCATCAAC             | TATTCCATGC   | TTAGCCTCGG    |
|-------|---------------|--------------------------|-------------|-----------------------|------------------------|------------------------------|-----------------------------------|------------------------|--|---------------|
|       |               | TITICTAGGAC              | GTCTCCTATC  | አርተ:አር፡ ያተርተ          | כנוגיו או יהאאכ        | TOTAGACIGAC                  | TRITTOCACCG                       | GCGGTAGTTO             | ATANGGTACG   | ANTICICACI    |
| 11401 |               | GCCCGCAAGA               | TATACCATAC  | CCCTTACGTT            | CUCATACACA             | ACCIONATION                  | CHACCHOOLD                        | TTCTACATGC             | OCATOCCCCT<br>CCTAPPCCCCA  | CAAGGTTGT '   |
|       |               | ראיסרפוורו               | MANGENTA    | GASAATTAL AA          |                        | 11.3. 11.5. 11.1             | Service Action                    |                        |  | CHENTERON     |
| 11501 | ACCTTOACCG A  | ACCACCTGGG<br>TOCTGGACCC | CCTTTATCCC  | MACCACACACA           | ACCAL AND SIG          | CC STEAN SCINE<br>GCACTCGCAC | אייכר היהיה אייכ<br>איימקב במכנים | COCTCGAGTC             | GCTGGCCCTC   | GACTACOTAST   |
| 11601 |               | GOCCCTGGCT               | GGCATGGGCA  | GCGGCGATAG            | AGAGGGGAAG             | TCCTACTTY                    | ACGCGGGCGC                        | TGACCTGCGC             | TGGGCCCCCAA  | GCCGACGCGC    |
|       |               | CCGGGACCGA               | CCGINCCCGT  | CGCCCCTATC            | TETECOGCTC             | AGGATGAAAC                   | TOCOCCCACG                        | ACTGGACGCG             | ACCCCCCCCTT  | caschacae:    |
| 11701 | CCTGGAGGCA G  | acteaaacca               | GACCTCAGGCT | GGCGGTGGCA            | בבבנונונונינוניניניני  | CTGGCAACGT                   | CGGCGGCGTG                        | GACCAATATO             | ACCAGGACGA   | TGAGTACGAG    |
|       | COACCTCCOT C  | CCACCCCGGC               | CTGGACCCGA  | CCGCCACCGT            | පහරපද  <br>            | GACCGTTGCA                   | GCGCCGCAC                         | CTCCTTATAC             | recreeroer<br>Pst  | ACTICATIGCTIC |
| 11801 | CCAGAGGACG G  | OCCAUTACTA               | AGCGGTGATG  | TITCIGATEA            | CATCATGCAA             | GACCCAACCC                   | ACCERREGART                       | оседосавса             | CTGCAGAGCC   | AGCCGTCCC*!   |
|       |               | CCCTCATGAT               | TEGECACTAC  | ANAGACTAGT            | CTACTACGTT             | Character                    | TOGGCCCCCCA                       | دودددوددود             | GACGICTCGG   | TCGGCAGGCr    |
| 11901 |               | ACGGACGACT               | OGCCCCAOCT  | CATGGACCGC            | ATCATCTCC              |                              | CAATCCTNIAC                       |                        | AGCAGCCGCA   | GOCCAACCT     |
|       | GGNATTGAGG 1  | TOCCTOCTGA               | CCCCCCTCCA  | GTACCTGGCG            | TAGTACAGCG             | ACTGACGCGC                   | GTTAGGACTG                        | CCCAMGGCCG             | regregaçõe   | COMPLIANCE    |
| •     |               |                          |             |                       |                        |                              |                                   |                        |  |               |
| 12001 |               | TTCTOGAAGC               |             |                       | ACCCCACGCA             |                              | CHOGCGATCG TANACGCGCT             | TAMACGCCCT             | GGCCGANAAC   | ACCCCCATC     |
|       | GAGAGGCCTT A  | AAGACCTTCG               | CCACCAGGGC  | CGCGCGCGTT            | TCCCCTCCCT             |                              | GACCCCTACC                        | GACCCCTACC ATTICCCCCGA | ccaacrifing  | TCCCGGTAG     |
| 12101 | OCCCCBACGA G  | GGCCGGCCTG               | GTCTACGACG  |                       | מכמכמוסטכב             | CCTTACAACA                   | GCGCCAACGT                        |                        | CITYGACCOGC  | TOGTOGGGA     |
|       | CCOORCIDET    | CCGGCCCGGAC              | CAGATICCTUC | GCGACGAAGT            | CCCCCCCCC              | GCANTISTANT                  | CGCCGTTGCA                        | ceremogras             | GACCTGGCCG   | MCCACCCCC"    |
| 12201 | TOTOCOCOAG G  | оссотовсес               | AGCGTGAGCG  | CCCCCAGCAG            | CAGGGCAACC             | TOCCCTCCAT                   | GGTTCCACTA                        | AACCCCTTCC             | TGAGTACACA   | OCCCCCCAN'    |
|       | ACACGCCCTC C  | CGCCACCGCG               | TCCCACTCCC  | GCGCGTCGTC            | GTCCCGTTGG             | ACCCCAAGGTA                  | CCAACGTGAT                        | TTGCGGAAGG             | ACTCATGTOF   | COGCICOGITA   |
| 12301 | 919000000     | GACAGGAGGA               | CTACACCAAC  | TPROTOTOGG            | CACTGCGGCT             | AATGGTGACT                   | GAGACACCGC                        | AAAGTGAGGT             | GTACCAGTCT   | GOCCCAGACT    |
|       | -             | CTGTCCTCCT               | GATGTGGTTG  | AMCACTCGC             | GTGACGCCGA             | TTACCACTGA                   | CICIGIOCOG                        | TITICACTOCA            | CATGGTCAGA   | CCCGGTCTGA    |
|       |               |                          | Pst         | _***                  |                        |                              |                                   |                        |  |               |
| 12401 | APPETETECA G  | GACCAGTAGA               |             | CANGGCCTGC AGACCGTAAA | CCTGAGCCAG             |                              | ACTINGCARGG                       | gcttcttoggg            | GIRCGOGCIC   | CCACAGGGGA    |
|       | TAMAMANGOT C  | CTGGTCATCT               | GTTCCGGACG  | TUTGGGATT             | GGACTCGGTC             | CGAMAGETET                   | TRANCGIECE                        | CGACACCCC              | CACGCCCGAG   | OCHCACGCT     |
| 12501 | CCGCGCGACC G  | GIGTCTAGCT               | TOCTGACGCC  | CAACTEGGGG            | CIKITIKATIKIC          |                              | GUNCETTUANG                       | GACAGTGGCA             | acatatecea   | CGACACATAC    |
|       | ODCCCCTGG C   | CACAGATCGA               | ACGACTISCGG | CITTGAGCGCG           | האכאאכלאכה             | ACGATTATCG                   | <b>ตลตลกกลาย</b>                  | CTGTCACCGT             | CCCACAGGGC   | CCTGTGTATG    |
| 12601 | CTAGGTCACT 1  | TOCTOACACT               | GTACCGCGAG  | GCCATAGRITC           | AGGCCCATCT             | CCACCAGCAT                   |                                   | AGATTACAAG             | TOTICAGCCGC  | טכפכידמפהקכ   |
|       |               | ACCACTGTGA               | CATGGGGGTC  | CGGTATCCAG            | TECHEGINGA             | CCTGCTCGTA                   | TYTANAGGTCC                       | TCTAATGTTC             | ACAGTCGGCG   | CGCGACCCCG    |
|       |               |                          |             |                       |                        |                              |                                   | **                     | Direction of the contract of t |               |
| 12701 | ADDAGGACAC C  | GOCCAGCCTG               | GAGGCAACTC  |                       | TAAACTACCT GCTCAACCAAC |                              |                                   | <b>GTTGCACAGT</b>      | TTANACAGOG   | ACCACCAGCG    |
|       | TCCTCCTGTG (  | CCCGTCGGAC               | CICCOTITION | ATTITICATICGA         | רהאכיוראיזויה          | CCCACCGTCT                   | TETARKAGAA                        | CAACGTGTCA             |  | TCCTCCT (*)   |
| 12801 |               | TACOTOCAGE               | AGAGCCTCAG  |                       |                        |                              |                                   | CTRICACATIGA           | CCGCGCGCAA   | CATCGAACCC    |
|       | GTANAACGCG A  | ATGCACGTCG               | TOTACACTO   | GGAATTGGAC            | TACTOCKICTOC           | CCCATTRCGG                   | מובנאנאנונננ                      | CACCTGTACT             | 0000000011   | GTACCTTCGGC   |
| ٠.    |               |                          |             |                       | =                      |                              |                                   |                        | -  |               |

# PMRKAd5gag MER6R2

|            |                       |              |                    | TITURE OF THE PARTY OF THE PART |                         | טיר אישריינירא דלי                      | טנינטננטנטעי                            | ACCCCCARGIA  | TTTCACCAAT    | CCTATICTINGA         |
|------------|-----------------------|--------------|--------------------|--|-------------------------|---|---|--------------|---------------|----------------------|
| 12901      | _                     |              |                    | AACC 14. 1. 1. AV  |                         |   |   |              | ANAGTOGITA    | C'XCTAGAAC'T         |
|            | CCCTACATAC            | GGAGTTTGGC ( |                    | TIVESTATION  | W. 1 1                  |   |   |              |               | Aracarcarcal T       |
| 13001      | ACCCUCACTO            | GCTACCGCCC ( | cctontrates.       |  | ATTCCAUCTC:             |   |   | CCACACACAC   |               | THERECACAA           |
|            | TECCETORC             | CGATGGCGGG   | GGACCAAAGA         | אינוטכנינכ   | TAAGGINGGAG GGAGINGGGAT |   | TGCTACC. IM                             |              |               |                      |
|            |                       |              |                    |  |                         |   | - 3                                     |              | •             | ;                    |
|            |                       |              | Harry Williams     |  | CACCAGGCAG              | ARRESTATION                             | GCGAAAGGAA AGCTTCCGCA                   |              | COCCMAGCAG    | CTTGTCCGA'I'         |
| 13101      | TICCCOCAN             |              |                    |  |                         |   | CGCTTTCCTT                              | TCGMOCCGT    | CCGCTTCGTC    | GNACAGGCT'.          |
|            | _                     |              |                    |  | 1 linefill              |   |   |              |               |                      |
| 1          |                       |              | TOURSE TOURS       |  | TTCCANSTT GATAGGGTCT    |   | CTTACCAGGA                              | CTCCCACCAC   | CCCCCCCCCCC   | CICCIONOC:           |
| 13201      | CTAGGCGCTG            | COCCCCCC     | Cheminate          |  | AAGGTTCCAA              |   |   | GAGCGTGGTG   | 0000000000    | CACCACCC             |
|            | מאורנפונמו            |              |                    |  |                         |   |   |              |               |                      |
|            | . ,                   |              | Auronn             |  | . JANAAAAA              | ניינינייניניני                          | CATTICCCM                               | CAACGGGATA   | GAGAGCCTAG    | TOGACANIA1           |
| 13301      | ADDAGGAGTA            | CCTANACAAC   | TCGCTGCTGC         | TCCTGCTGCTGC WALLACTURE  |                         | מערמטענונכ                              | GTAMGGGTT                               | GTTGCCCTAT   | CTCTCGGATC    | ACCTI:TTCTA          |
|            | recreerest            | CONTINUENC   | AGCGACGACG         | שרפוירני ווינייר   | 00111110                |   | AV.                                     | AGCCACGACC   | GTCAGCGGGG    | rendendian           |
| 13401      | GAGTAGATGG            | ANGACOTACG   | CCCACCACCA         | CACCOACTA  | ברשאירו ניסר            | מייינית בכייב                           |   | TOUGHELINGS  | CAGTCGCCCC    | AGACCACACT:          |
|            | CTCATCTACC            | TICTGCATGC   | COLCUICOL          | GICCCICCAC   | ואניתו ירונים איכני     | CIARACIA                                | 100000000000000000000000000000000000000 | 00100101     |               | a monthimetry A.A.A. |
| 13501      | GAGGACGATG            | ACTOGGCAGA   | CONCARCACE         | GTCCTGGATT   | TEGENORANG              | TYGGAACCCG                              | TTICCCICACC                             | TTCGCCCAG    | GCT GGGGAGA   | TACABABITT           |
| )<br> <br> | CICCIGCTAC            | TOAGCCGTCT   | CENTERING          | CAGGACCTAA   | ACCCTCCK:TC             | ACCOMICGGC                              | MARCINCOTTIG                            | AAIXCGGGGIC  | רושררורוי     |                      |
| 13601      |                       | GCATGATGCA   | AAATAAAAA          | CTCACCAAGG   | CCATTACACC              | CAGCGTTGGT                              | TITICITISTAT                            | TCCCCTTAGE   | ATGCGGCGCG    | COCCONTGTA           |
| 10001      | and the second second | CONTRACTOR   | THINTIE            | GAGTGGTTCC   | GGTACCCOTCG             | CTCCCAACCA                              | ANAGAACATA                              | ACCCCANTCA   | TACCCCCCCC    | gccact near          |
|            |                       | כפושרושרפו   |                    |  |                         |   | CATTROCETECT                            | CCCTTCGATG   | CTCCCCTGGA    | CCCGCCGTTT           |
| 13701      | TCACCAACOT            | CCHCCACCCT   | CCTACCIACIAO       | יוניוניניונינין.   | CALLAR CR. C. M.        | שניניניניניניניני                       |   | GGGAAGCTAC   | GAGGGGACCT    | GGCCGCCAA            |
|            | ACTCCTTCCA            | CCACCACCCA   | GGATGCTCTC         | ACACCACICO   | רפרוניניניני            | Week Constant                           |   |              |               |                      |
|            |                       | Kon          |                    |  |                         |   |   |              |               | CHGGTGGACA           |
| 13801      | OTCICCTCCGC           | GGTACCTGCG   | <b>GCC_TACCGGG</b> | GRANGANACA   |                         |   | GCACCCC INI                             | SCHOOL STATE | CCCACACATA    | GALLACTE             |
|            | CACCOAGGCG            | CCATOGACGC   | CCCATCCCCC         | CCCTCTTIGE   | CCTAGGCAAT              |   | CCTCCCIAIA                              | ACTIVITIES   | on concurrent | Contractions         |
| 11901      | ACAAGTCAAC            | GGATGTGGCA   | TCCCTFAACT         | ACCAGAGAACGA   | CCACACCAAC              | TITCTCACCA                              | CCCTCATTCA                              | ANACANTGAC   | TACAGCCCGG    | Greater Contract     |
|            | TOTTCAGTTG            |              |                    | TAGICTIVICT  | <b>AGTOTOGTTO</b>       | ANAGACTEGET                             | CCCAGTAAGT                              | THEFTACIO    | ATGICTAGICC   | ררוניפוני            |
| 14001      | CACACAGAGA            | ATCANTETTO   | ACCACCGGTC         | GCACTGGGGGC  | CKICCACCTICA            |   | GCATACCAAC                              | ATGCCAAATG   | TOAACGAGTT    | CARTTERIC            |
| 700        | CENTRALITY            | TAGTTAGAAC   |                    |  | CCCCCTCGACT             | TITCGTACK                               | てらすんすらにすずら                              | TACCCTTTAC   | ACTIGCTCAA    | GTACTOWNIA           |
| 14101      | BATAACTTA             | AGRICATIONS  | _                  | COCTTGCCTA   | CTAAGGACAA              | Transmode                               |   |              |               | CCCGWRRCA            |
| 10151      | THATHCAAAT            |              |                    |  | GATTCCTVTT              | אהזבכאנז־זכ                             | GACTITIATIC                             | TCACCCACCT   | CAAGTGCGAC    | פמיבוררניו           |
|            |                       |              |                    |  | Pvril                   |   |   | •            |               |                      |
| 1014       |                       | CACTATARE    | ATAGACCTTA         | TGAACAACGC   | GATCHTGGAG              | CACTACTTGA                              |   |              |               |                      |
| 10951      |                       |              |                    |  | CTAGGACCTC              | CITCATGAACT                             | THEACCEGIC                              | TOTOTINGCCC  | CAAGACCTTT    |                      |
| , ,        |                       |              |                    | _  |                         | CHETTEREAT                              | CCCTCCCCTA                              |              |               |                      |
| 14201      |                       |              |                    |  |                         | にんにんんにんのする                              | てらられていている                               | ATATOTTEG    |               |                      |
|            |                       |              |                    |  |                         | כבוגונינבונינים                         | ATCUCANAGE                              | GGCAACCCTT   | CCAGGAGGGC    | TTTAGGATCA           |
| 14401      |                       |              |                    |  |                         | CANTANTO                                |   |              | GGTCCTCCCG    | ANATCCTAGT           |
|            | TAMAAACGACG           | S GTCCTACGCC | CCACCTGANG         | TOTAL CITY OF  |                         | ייייייייייייייייייייייייייייייייייייייי |   |              |               |                      |

Figure 15I

| THE COLUMN TO THE COLUMN THE COLU | ACCAT<br>ACCAT<br>ACCAT<br>ACCATOR   | CCAAGA<br>GGTTCT<br>Ascl                  | בסכנו.                                      | NGCOA<br>ICÓC'I<br>NGCAA<br>ICGIT'I                   | באביבי.<br>הייניאי:                    | MCCT<br>GAAGA<br>CTTCT<br>AAGCGG                        | CCATT<br>GGTAA<br>CTAGA  | ATCCT.<br>TACCTA                       |
|--|--|---|---|---|--|---|--|--|
| CCCCACCGC TCCCATCGC ACCCATTCGC ACCGTAAGCG GAGAAGCCTC CTCTTCGGAA  | GCAGCTGCTTA<br>CGTCGACCAT  |   |   |   |  | CCACCACCT AAAATGAAGA TTTTACTTCT GCACCTAGGG              |  |  |
| CHTCTCCGC TCAACAGATCA ACTTCCTAGT ACCCGAGGTC TGGGCTCAGT   | TACANCETAA TANGCAATOA CAGCACETTE ACCEAGTACE ATGTTCATEG TGGTCATEG TAGTTCATEG TAGTTCATEG TAGTTCATEG TAGTTCATEG TAGTTCATEG TAGTTCATTCATEG TAGTTCATEG TAGTTCAT | OTTOCCOGTO<br>CAACOGGCAC                  |   | -   |  | GGTAGCTAGG<br>GCGCTATGCT<br>CGCGATACGA<br>CGCGCACGTC    | CCGCACCACAGC GGCGTCGTCGTCG CCGCCCCCCA  |  |
| AGATGACACC<br>GTTACTGTGG<br>GTTGAAGGACA<br>CACCTCCTGT<br>CCCCTGGGCA<br>GGCGACGCGT  | TAMBEANTON CAGCACCTTC ATTCGTTACT GTCGTGGAAG TCCTGACCTA ACCTGGGGCT  |   | AGCGAAAGGG                                  |   |  | CAGCTACTIGC<br>GCGAAGCCCG<br>CGCCTCGGGC<br>CCTGCTTAAC   |  |  |
| CCTCGACTTGAA<br>GCTCGAACTT<br>AATGCAGCG<br>TTACGTCGGC<br>GCTYGCGCCTC   | TAAGCAATGA<br>ATTCGTTACT<br>TCCTGACGTA   |   | ACCTIGATIONA<br>TOCACAGGIT                  | CACCCTACCG<br>CTYCCGATCGC<br>GTCTCCCCGC<br>CAGACCGCCG |  | CCCCTCGTCGTCC ACCGTCGTCACC TCCCACCACCG CCACCGCCCCCCCCCC |  |  |
| CCSANCATOC<br>CCSANCATOC<br>CACCACTOCC<br>CACCACTOCAC<br>ACCACTOCATO   |  |   | TCTCTGACCC<br>AGAGACTGGG                    | CAGATCACGG<br>GTCTAGTGCC<br>CCTGGGCATA<br>GGACCCGTAT  |  |   | TECCECCOS ACTOR AC | COCCOCCC                               |
| GGANGIGAC<br>GCTACACTTG<br>TCCAACGTTG<br>AGGTTGCGC<br>TCCGGCTTCG   |  | CGCGCGCTCTA                               | AACTCATCIG CCAGTTTACC TTGAGTAGGC GGTCAAATGG | CCTOCTOTO<br>GGACGAGAGT<br>TTTACAAGAC                 | המאכדהה<br>בככסהארה<br>באכאארהכה       |   | -  | TCCAGCGGG<br>AGGTCGCCGG                |
| CCCCACTYTT<br>GGCGTTACAA<br>GAAAGACAAC<br>CCTTCTTG<br>AACCCCCAC  | APPACAGGAA<br>TCCTGTCGTT<br>CCGAATCCGC   | OCCTTACACO<br>TTCCCCTCCA<br>AAGGCGAGGT    | AACTCATCCG<br>TTGAGTAGGC                    | TGAAAACGTT<br>ACTTTTGCAA<br>TGCCCCTACG<br>ACGGGGATGC  | ACACAGGCTG<br>TGTGTCCGAC<br>CTGGGGCGCG | GACCCCGCGC<br>CCAGTCTCCA<br>GGTCACAGGT<br>GCCGCCGACC    | GOCCOCCAGO<br>CCCGCCCCA<br>CCCGCCCCAA  | CANCATACAT CANCATACAT                  |
| GGTAACATTC<br>CCATTGTTGG<br>GGAGGCGG<br>CGTGGCCGG<br>GGCTGACGAG  | CCCCTTACAG<br>GGGGACTGTC<br>ACCCTCAGAC   | TOGOGENETIO<br>GEOGENETIOS<br>GEOGENETIOS | GTCFACTCCC                                  | CCACCOTICAG<br>GGTGGCAGTC<br>ACGCCGCACC<br>TGCGGCGTGG | CCCACCAATA<br>GGGTCGTTAT<br>ACCGCGCGCC | TOCCOCCCC<br>CACCCCCCCA<br>GTCCCCCCCT<br>COTCCCCACC     | GCAGCGGTGG<br>CTCGAAGGCT<br>GAGCTTCCGA<br>CAGGGGCAAC   | GACTCGTAGT CTGAGCATGA                  |
| TCTGGAGGGT<br>ADACCTCGCA<br>AACAGCAGTG<br>TTGTCGTCAC<br>ACCGTCTGC  | OCTONICAAA<br>CCACTAGIITI<br>AACTACGCG   | TTGATGCGGC<br>TGATGCAAGA<br>ACTACGTTCT    | CCACCAGOCC<br>GCTCOTCCGG                    | CCCACCATCA<br>OCCTOCTAGT<br>CTCACCCCAG<br>OACTGCGGTC  | CCTTATATCO<br>GGAATATAGC<br>CGCGGGCACT | ACTACACGCC<br>TCATGTGCGG<br>GCGCGTAGCA                  | COCGCATCGT<br>SIII<br>ATGCGGGCG<br>TACGCCCGGC<br>CTCAGGGTCG  | GAGICCCAGC<br>AAACTACTTA<br>TITGATGAAT |
| CCTACGATGA<br>GGNTGCTACT<br>AGGGGCAGC<br>TCCGCGTCG<br>GGCGACACCT   | AGAAGAACC<br>TCTTCTFTGG<br>Kpm<br>CCTTGCATAC   | GGAACGTATG<br>TTGCCAGACA<br>AACGGTCTGT    | OCTTCTACAA<br>CGAAGATUTT<br>Asci            | CCCOCCAGCC<br>GGGCOOTCGG<br>GTGACCATTA<br>CACTGGTAAT  | GCATGTCCAT<br>CGTACAGGTA<br>AGTGCGCGTG | TCACGCGCAC<br>GAGGCGCGCA<br>CTCCGCGCGT<br>GACGGCGGAG    | ACCCCCCCC TCCCCCCCCC TCCCCCCCCCCCCCCCCC  | TCACGATACT TTGCAAGAAA AACGTTCTTT       |
| 14501<br>14601<br>14701  | 14801  | 15001                                     | 15101                                       | 15201   | 15401                                  | 15601   | 15801  | 16001                                  |

Figure 15J

# PMRKAdSqag MER682

| ,     |                 |  |                                  | 2440442000                                  |                   | ATTACAARCC (               | CCCAAACCTA A                    | THE PROPERTY OF THE PROPERTY OF THE PROPERTY ANGEGOSTER ANARGMAIN | MARRIAN AND AND AND AND AND AND AND AND AND A | GNACIATION        |  |
|-------|-----------------|--|----------------------------------|---|-------------------|----------------------------|---------------------------------|---|---|-------------------|--|
| 6101  | CCAGGTCATC      | CCCCCCACA                                | GCGCCGGGG TCTATCG C. CCCGGAAAAAA | GGCTTCTTC                                   | C.T.Tr. Trythric  | TAMINITACIO GOCTITICGAT    | GCCTTTCGAT '                    | TITEGECECAGT 1  | THICHILL                                      | CTTTCTACT.        |  |
|       |                 |  |                                  |   |                   |                            | •                               | Sall  |   |                   |  |
|       |                 |  |                                  | ATOMOGRAPH CHARLES CHARLES CHARLES CHARLES  | נו אבוניניניניעני |                            | CACTCGAAAG                      | COOT  | AAAACGTGTT                                    | MCCGACCC.         |  |
| 6201  | GATGATGAAC      | TIGACGACUA                               | GENERALINE IG                    | CHARACTER CONCURS. CONCURS.                 |                   |                            |                                 |   | TTTTGCACAA                                    | AACGCTGGra .      |  |
|       | CTACTACTIO      | AACTGCTGCT                               | רכאררונאר                        | משטישים איני                                |                   |                            |                                 |   | GACCTGCTTG                                    | NOCAGGGGAA        |  |
| 6301  | GCACCACCOT      | AGICTITACO                               | CCCGTNAGC                        | וכעכעונערונפ                                |                   |                            |                                 |   | CTGGACGAAC                                    | TCGTCCGGTT        |  |
| ,     | Corocioaca      | TCAGAMIGE                                |                                  |   |                   |                            |                                 | ANCCCANCAC (  | CTAGCCTAAA                                    | <b>GCCCGTAACA</b> |  |
| 6401  | COMOCOCCIIC     | CACCACACACACACACACACACACACACACACACACACAC | CCTACGGMA                        |   |                   |                            |                                 |   | GATCOGATTT                                    | COCCCATTGT        |  |
|       | GC-ICCCOCAG     | רררורשער                                 |                                  |   |                   |                            |                                 |   |   | Kini              |  |
| ,     |                 | •  |                                  | W.C. B. |                   | AAAGCGAG                   | TETOGTISACT                     | TOCACCCAC (   |   | ATCOURACCE        |  |
| 16501 | CHECAGEMON      | Arghrigger o                             |                                  | AGGCTTCTTT                                  |                   |                            |                                 | ACCGTGGGTG (  |   | TACCATVAGGT       |  |
| 5601  | Margine Agencia |  |                                  |   | GGAACCTEAN        | CHAGAGCCCG                 |                                 |   |   | COCCGGGACT        |  |
|       | TCGCGGTCGC      | _  |                                  | TTTACTGGCA                                  | CCTTGGACTC        | CACCTCGGGC                 |                                 |   |   | OCCESCO TO        |  |
| 16701 | GERCOTOCAG      |  | TTCAGATACC                       | CACTACCAGE                                  | AGCACCAMTA        | THECCACCEC                 | CACAGAGGGC                      |   | ANACOTOCOC                                    | GGTTGCCTCA        |  |
|       | CCCOCACOTC      | _  | _                                | CTCATGGTCA                                  | TCCTRGTCAT        | AACGGTGGCG                 | grencheed                       | -   | TTTCAGGGG                                     | CCAACGGAGT        |  |
| 16801 | geograpeag      | ATOCCOCOGE                               | GCAGGCGGTC                       | -   | CGTCCAAGAC        | CTCTACGGAG                 | GTCCAAACTG                      |   | Grancacerr                                    | TCAGCCCCCC        |  |
|       | COCCACCOCC      | TACOOCOCCA                               | COTCCGCCAG                       | CGACGCCCCC                                  | GCAGGTTICTG       | GAGATGCCTC                 | CACGITITIACC                    |   | CANAGEGERA                                    | ACT CONTOUR       |  |
| 16901 | 000000000       | CCGTTCGAGG                               | MAGTACGGCG                       | CCCCCAGCGC                                  | GCTACTGCCC        | GNATATGCCC                 | TACATOCITIC                     |   | ACCUCURACT                                    | ATCOTOCCT         |  |
|       | CCGCGGGCGC      | GOCAAGCTCC                               | TTCATGCCGC                       | CCCCCICCCC                                  | CCATGACOO         | CTTATACGGG                 | ATCTAGGAAG                      |   | Transported                                   | TWOCACCOM         |  |
| 17001 | CACCTACCOC      | CCCAGAAGAC                               | GAGCAACTAC                       | CCGACGCCGA                                  | ACCACCACTG        |                            | CCCACATATACGC                   |   | CCGTGCTTAGC                                   | CCCGATTICC        |  |
|       | OTOGATOGCO      | _  | CTCGTTGATO                       | <b>OCCTGCGGCT</b>                           | TROTOCTOAC        | CTTGGGCGGC                 | GCCCCCAGCG                      |   | מכשרפעררם                                     | COCC INDEXIO      |  |
| 17101 | GRECOCAGGO      | -  | ACCAGG CAGG                      | ACCCTOGOTOC                                 | TRECOVICARE       |                            | CCCAGCATCG                      |   | GGICTITIGIO                                   | OTTCTTICCAG       |  |
|       | CACGCGTCCC      | : ACCGAGCGCT                             | recressive                       | TEGGACCACG                                  | ACCIGITAGACG      | CACGATGGTG                 | GGCTCGTAGC                      | AAATTTTCGO  | CLAGAMACAL                                    | Sphi              |  |
|       |                 |  | - تعلقائد ململمالان              |   | APPLICACE         |                            | AGANTGCACC GTAGGAGGG CANGGCGGGC | CATGGCCGGC  | CACOOCCITOA                                   | COCCURRENT        |  |
| 17201 | TATACCICCI      | GRECTION OF                              |                                  |   |                   |                            | CATCCTCCCC                      | GTACCOGCCG  | GTGCCGGACT                                    | acccaccaty        |  |
|       | Ē               |  |                                  |   | Sphi              |                            |                                 |   |   |                   |  |
| 17301 | aconcoraca      |  |                                  |   |                   | CCCCTATCCT                 | GCCCCTCTT                       | TAAGGTAACT  | Ancidecece                                    |                   |  |
|       | COCNOCNOC       | : orogradice                             |                                  |   |                   |                            |                                 | ATCABABTAB  | AAAGTCTGGA                                    | CTCTCACGUT        |  |
| 17401 | GTOCCCOGA       |  |                                  |   | ACTICATTION       | THE PLANTS                 |                                 | TAGTITIAIL  | TITCAGACCT                                    | GAGAGTECETA       |  |
|       | CACOOCCT        | F AACGTAGGCA                             | A CCCGGAACGIC                    | בפרפורורופ                                  |                   |                            |                                 |   |   | Ecofil            |  |
|       |                 | THE STATE OF STATE                       | P THETAGAATE                     | GAAGACATCA                                  |                   | אכידידופניוב יונידפתיכנינס | CGACACGGCT                      | CGACACGCCT CGCCCCCTT CATGGGAAAC                                   | CATGGGAAAC                                    |                   |  |
| 105/1 | GCGNACCAGG      |  |                                  |   |                   | TGAAACGCAG AGACCGTGAC      | GCTGTGCCGA                      | GCTGTGCCGA GCGCGGGCAA   | GTACCCTTTG                                    | ACCOLTCTA"        |  |

Figure 15K

| ;<br>;             |              |              |                          |             |               |                          |             |                   |              |  |
|--------------------|--------------|--------------|--------------------------|-------------|---------------|--------------------------|-------------|-------------------|--------------|--|
| EcoRV              |              |              |                          |             |               |                          |             |                   |              |  |
| TCGGCACCAG         | CAATATGAGE   | CCACCGCCCA   | TCAGCTGGGG<br>AGTGGAGGCC | CHUCULULUS  | ACCCCTAAT     | AAAATTTAG<br>TTTTAAAAGCC | ANGETOCCAA  | THETHOATAC        | CGTCGTTCC::  |  |
| CTOGAACAGC         | AGCACAGGCC   | AGATGCTGAG   | GGATAAGTTG               | AAAGAGCAAA  | ATTRICANCA    | AAACCACCAT               | GATGCCCTGG  | CCTCTGGCAT        | TAGCCCCCA!   |  |
| Succi i Succi      | •            |              |                          |             |               |                          |             | -                 |              |  |
| GTOGACCTOG         | CCAACCAGGC   | AGTGCAAAAT   | AAGATTAACA               |             |               | CCCOTTAGAGG              |             | OCCCC+CCAO        | ACAGTGTCT"   |  |
| CACCTGGACC         | GOTTROTTCG   | TCACGITITA   | TTCTAATTGT               | CATTRUCAACT |               | PARTICINE.               |             | CCGGCACCIC        | ופורארארייי  |  |
| CAGAGGGGCG         | -            |              | CCGACAGGGA               | NGAMETERS   | Christicicity | TACACCACC                | TCCCTCGTAC  | CHECHECGIO        | ATTITICGITIC |  |
| Grenecede          | •            |              | GRETGICCE                |             |               | COTTABOOTING             | Christian C | CCCCCGCCGA        | CACCCAGCA    |  |
| CCTOCCCACC         | Transcendent |              | AGCGCGGTA CCGATGGCT      |             |               | GCATTGCGAC               |             | GGGGCCGCT         | GINGGICCTI:  |  |
| our reconstruction | -            |              |                          |             |               |                          |             | Prof              |              |  |
| AAACCTOTOC         | TOCCAGOCCC   | GACCGCCGTT   | GTTGTAACCC               | GTCCTAGCCG  | COCGIOCOTO    | دوددودوددو               | CCAGCIAGICC | OCGATOSTIO        | COCCCOTAG    |  |
| TTTOGACACO         | _            |              | CAACATTICAG              | CAGGATCGGC  | מכפכשפעפעכ    | ට්ට්ට්ට්ට්ට්ට්ට්         | GOTCGCCAGG  | CGCTAGCAAC        | GCCTAGGCATC  |  |
| CLACTORICAA        | _            |              | CCATCCTGGG               | TCTGGGGGTG  | CAATCCCTVA    | ACCCCCACG                | ATCCTTCTOA  | TAGCTAACGT        | greatatona   |  |
| GETCACCOTT         | _            |              | CGTAGCACCC               |             | GITTAGGGACT   | TCCCGCCTGC               | TACGAAGACT  | ATCGATTGCA        | CACCATACA    |  |
| TATTE ATTENT       | •            |              | GCAGCTGCTG               | ACCECCEC    | CACCAGATT     | CCAAGATYBC               | TACCCCTTCO  | ATGATGCCGC        | ACTOCITION A |  |
| ACAGTACATA         | _            | _            | CCTCGACGAC               | TCGGCGGCGC  | GCCCCCCCAAA   | GGTTCTACCG               | ATGGGGAAGC  | TACTACGOCO        | TCACCAGA: P  |  |
| CATATATATA         | _            |              | GTACCTCAGC               | CCCGGGCTGG  | TCSCAFFTTRIC  | ככפכפככענכ               | GAGACOTACT  | TCAGCCTGAA        | TANCAAGTIT   |  |
| GTACGTGTAG         | _            |              | CATGGACTCG               | Greecedance | ACCTCAAACG    | CCCCCCCTCC               | CTCTGCATGA  | AGTCGGACTT        | ATTOTTCAAA   |  |
| ACABBUCULA         | _            |              | GTCACCACAG               | ACCRETCCCA  | GCGTTTGACG    | CTRECHETICA              | TCCCTGTGGA  | CCOTGAGGAT        | ACTRICUTALLY |  |
| TCTTTGGGGT         |              |              | CACTOGRAPE               | TOGCCAGACT  | COCMACTIC     | GACCCCAAGT               | AGGGACACCT  | <b>GCCACTCCTA</b> | TCACCICATGA  |  |
| COTACAROGC         | _            | CTAGCTV;TGG  | GTGATAACCG               | TGTGCTGGAC  | ATGGCTTCCA    | CCTACTITICA              | CATCCGCGGC  | GTGCTGGACA        | OCCCCCTA     |  |
| GCATGTTCCG         | _            | GATCGACACC   | CACTATTAGE               | ACACGACCTG  | TACCGAAGGT    | GCATGAAACT               | GTAGGCGCCG  | CACCACCTGT        | CCCCCCCATA   |  |
| THITAAGECE         | -            |              | COCCCITORICT             | CCCMGRATG   | CCCCAAATCC    | TTGCGAATGG               | GATCAAGCTO  | CTACTCCTCT        | TGAANTAAN    |  |
| AMANTICGGG         |              |              | CCCCCACCCA               | CCCTTCCCAC  | GGGGTTTAGG    | AACGCTTACC               | CTACTICGAC  | GATCACCACA        | ACTTTATTT:   |  |
| CTAGAAGAAG         | * ADDACCATCA | CAACGAAGAC   | GAAGTACACG               | ACCANGETOA  | GUNGUNAAA     | ACTCACGTAT               | TTCCCCAGGC  | CCCTTAITICT       | GCTATAAATA   |  |
| GATCTTCTTC         | -            | GTIOCTICTO   | CTTCATCTGC               | TCGTTCCACT  | CCTCGTTTTT    | TGAGTGCA1'A              | AACCCGTCCG  | CCCNATAAGA        | CCATATTIAL   |  |
| TTACABACTER        |              | ATAGGIGICG   | AAGGTCAAG                | ACCTANATAT  | GCCGATAAAA    | CATITICAACC              | TGAACCTCAA  | ATACCAGNAT        | CTCAGTGGTA   |  |
| AATGTTTCCF         |              |              |                          |             | COCCTATTIT    | GTANAGTTGG               | ACTIGGAGIT  | TATCCTCTTA        | GAGTCACCAT   |  |
| CCIAAACAGAA        | _            |              | AGTCCTANAA               | AAGACTACCC  | CANTGANACC    | ATTACAGE                 | TCATATGCAA  | AACCCACAAA        | TCAAAATC3C?  |  |
| GCTTTGTCTT         |              |              | TCAGGATTTT               | TICTGATOGG  | CITTACITATION | TACAATGCCA               | ACTATACGTT  | TIGGGTGTTT        | ACT TITACC'  |  |
| GGCCAAGGCA         | _            | GCANCANAT    | GGAAAAGCTAG              | AAACTCAACT  | GGAAATGCAA    | TITITICICA               | CTACTGAGGC  | AGCCGCAGGC        | AATGGTGAT    |  |
| cccorrect          | F AAGAACATTT | CGTTGTTTTA   | CCTTTCKIATC              | TTTCACTTY'A | にじずずがんらずず     | MANAGAGTIT               | GATGACTCCO  |                   | TTACCACTAL   |  |
| ACTIGACTOC         | -            |              |                          |             |               |                          | CATGCCCACT  | ATTAACAGAAG       | OTANCTCAC!   |  |
| TGAACTGAGG         | 3 ATTTCACCAT | . AACATGTCAC | TTCTACATCT               | ATARCTTAGG  | GCTCTGTGAG    | TATAMGAAT                | GTACGGGTGA  |                   |              |  |
|                    |              |              |                          |             |               |                          |             |                   |              |  |

Figure 15L

| 19301 | AGAACTAATO               | GGCCAACAAT                   | CTATCCCCAA                | CAGGCCTAAT                  | TACAPPGETT                               | TENGGGACAA 1              | TTTTATTGGT C                  | CTANTCTATT A | _                               | GGGTAATATG<br>CCCATTATAC         |
|-------|--------------------------|------------------------------|---------------------------|-----------------------------|--|---------------------------|-------------------------------|--------------|---------------------------------|----------------------------------|
| 19401 | GRIGITATIO               |                              | ATCGCAGTTC                | AATVICTIVITYE<br>TPACGACAAC | TAGATTTGCA                               | AGACAGAAAC A              | ACAGARCTITI C<br>TOTCTCGAAA C | CATACCAGCT T | AACGAACTA                       | TCCATTGG173<br>AGGTAACCAT        |
| 19561 | ATAGAACCAG               | . •                          | ATCHCGAATC                | ACCCACACT                   |  | CCAGATGITTA C             | CTTANTARCT 1                  | AAATCATGGA A | ACTGAAGATG<br>TGACTTKCTAC       | AACTI'ICCAAA<br>TTGAA' 2 :'''''' |
| 1961  | TTACTOCTTT               |                              | GTGTGATTAA                | TACAGAGACT                  |  |                           | AACAGGTCAG C                  | CHITTACCIA C | GOGANNAGA                       | TCCTACAGAA<br>ACGATGTCTT         |
| 19701 | TTTTCAGATA               |                              | AAGAGTTCAGA               | AATAATTTTG                  | CCATCGAAAT                               | CANTCTAMAT (              | COGITICANO C                  | CCTCTTTAAA C | CCTUTACTCC                      | AACATAGCOOT                      |
| 19801 | TOTATTECC<br>ACATAAACGO  |                              |                           | CTTCCAACGT                  | AAAAATTTET<br>TTTTTAAAGA                 | CTATTGGGTT '              | ACACCTACCIA ( TGTCGATGCT (    |              | ANGCCAGTOO<br>TICGCTCACC        | TOGCTCCC():                      |
| 19901 | OCTAGNOGAC               |                              | ACCTTGGARC<br>TGGAACCTCG  | ACCCTAGNCC TCC TCCCACC      | CTTGACTATA                               | TOTACAACGT ACCTOTACT      | CAACCCATTE A                  |              | CONTRACTOR                      | GGACGCGATA                       |
| 20001 | CCCTCAATGT               |                              | -                         | GTGCCCTTCC                  | ACATECAGET<br>TREAGGECCA                 | COGAGICTIC                | TTCTTTGCCA .                  | TTANANACCT ( | CCTTCTCCTG                      | CCGGCTCAT<br>GGCCCGAGTA          |
| 20101 | ACACCTACOA               | A OTCOAACTTC                 | AGGAAGGATG<br>TECTTECTAG  | TTAACATGGT                  | MANAGE PSII MANAGE TECTIFETE AGACETECTEG | TCCCTAGGAA                | -                             |              | <b>GCCAGCATTA</b><br>CGGTCGTAAT | AGTITICATA:<br>TCANACTAIN:       |
| 20201 | CATTHGCCTH               |                              | AGAAGGGGTA                | GGCCCACAAC                  | ACCOCCTCCA<br>TCX CCIGAGGT               |                           |                               |              | ACGACCAGIC<br>TGCTGGTCAG        | GAANTTGC'I :                     |
| 20301 | TATCTCTCCC               | - ·                          | GCTCTACCCT                | ATACCCGCCA<br>TATGGGCGGT    | ACCCTACCAA<br>TOCCATGGTT                 | CGTGCCCATA                | TCCATCCCT                     | -            | CCCCCCAAG                       | GCGCCGACCC                       |
| 20401 | CCTTCACGCC               | m ti                         | AAGGAAAACCC<br>TTCCTTTGGG | CATCACTGGG                  | CTCGGGCTAC<br>GAGCCCGATG                 | CTCCCTTATT                | ACACCTACTC                    |              | CCCTACCTAG                      | ATGGAACE FF                      |
| 20501 | TTACCTTCAAC              | D 13                         | AGANGGITGGC<br>TCTTCCACCG | CATTACCTTT                  | GACTCTTCTG<br>CTGAGAAGAC                 | TCARCTRICC<br>AGTERACEOG  | TCCCAATCAC<br>ACCGTTACTG      |              | CCCCCAACGA                      | CMMCTTTAA                        |
| 20601 | AAGCGCTCAC<br>TTCGCGAGTC | m ti                         |                           | GTTGCCCAGT<br>CAACGGGTCA    | GTAACATGAC<br>CATTGTACTG                 | CAMAGACTOS                | TTCCTFGTAC<br>AAGGACCATG      |              | TAACTATAAC<br>ATTGATATTG        | ATTK# ICTACC<br>TAACCGATGG       |
| 20701 | AGGGCTTCT/<br>TCCCGAAGA  | A TATCCCAGAG<br>T ATAGGGTCTC | ACCTACAAGG                |                             |  | AGAMACTICC                | AGCCCATGAG                    |              | CACCTACTAT                      | CATTTATGT                        |
| 20801 | CCTGATOGT                | A CAGOTOGGCA<br>T GTCCACCCGT | AGGATGTGGT                | ACACAACAAC<br>TGTGTTGTTG    | TCTGGATTTG<br>AGACCTAAAC                 | TTGGCTACCT                | ACCCCCACC<br>ACCCCCCACC       | TACGCGCTTC   | CTGTCCGGAT                      | GGGACGA1TG                       |
| 20901 |                          | C COCTTATAGG                 | CAAGACCGCA                | GTTGACAGCA<br>CAACTGTCGT    | TTACCCAGAA                               | AAAGETTECFT<br>TTTCAAAGAA | TYCCGATCCCA                   | CCCTTTGGCG   | CATCCCATTC                      | TCCAGTAACT                       |

Figure ISM

|       | <b>K</b> F.   |   | }ħ         | Ą.         | à R                      | N F.                     | <i>;</i> ; <b>t</b> . |            |            | <b>i</b> . "                               | S ±                      | £ 5                      | s t  | R is   | <b>B</b> H               | H B                      | £ 5  | e n  |
|-------|---|---|------------|------------|--------------------------|--------------------------|-----------------------|------------|------------|--|--------------------------|--------------------------|--|--|--------------------------|--------------------------|--|--|
| ,     | CCATOGACC   | CACGCCCTT :   | TCANAGATET | AUTTICTAGA | GOCCIGGTORS :            | COACTICAAGE              | AAAGCGTAC .           | CACCATONA  | GTCGTACTT  | OAGCGCCAC"<br>CTCGCGGTG                    | CTTTCAATAA<br>GAAAGTTAT" | CGCATCGCTA<br>GCGTAGCGAY | TCACTCCACA.                                    | CTATUTOTOTO  | GTTGCTCAGG<br>CAACGAGTCC | TGACCGTGCC<br>ACTGGCACGG | CGFAAGACTT   | GFFCFFCAC  |
| Ramit | CTCC  | TGTACCTGCG  | AAAGCCATTG | TTTCOGTANC | TAGTCAATAC               | PTCTCACCAG<br>AAGACTGGTC | AACGCTGGAA AAGTCCACCC |            | TAGTGTTOOG | CAGCTTCCTO                                 | ACTAGACACA<br>TOATCICTOT | OCTTOTOCCO<br>CCAMBACOCC | GGTGAAGTTT<br>CCACTTCAAA                       | COCCAGITOC   | GGTCCTCCGC               | CATCAAAAGG               | AAGAACATGC<br>TTCTTGTACG                           | TRAFTRACCE ANCAGACECE GREGIVICACE CAGCACETTO CERCAGATIT GOAGATATITE ACCACATITE GOCCCCACE SITICTICACA ACTIANCEGEC CIGITECEGEC CAGCACATITE GOCCCCACEG SITICTICACA ACTIANCEGEC CIGITETAAAG CCGGGTUGC CAGAAATTEC |
|       | CATCACTTITT<br>GTACTGAAAA   | ATCGNANCCG<br>TAGCTTTGGC  | GCAGGAACTG | COTCCTTGAC | GCCTGCGCCA<br>CGGACGCGGT | CCTTTGGCTT               |                       |            | TCAGGGTACC | AACAGCTCTA                                 | AAAATAATGT<br>TTTTATTACA | ANTCANAGO                | ACAACCATCC GCGGCAGCTC<br>TOTTGGTAGG CGCCGTCGAG | TERRESCENCE GEOCITACEO COCOMOTIVO<br>ACCCCIGADO COCOMOCO GEOCITAMO | TCCGCGTCCA<br>AGGCGCAGGT | ACCOTAGTOD<br>TOCCATCACC | GCCTTCAGAG<br>CGGAAG II: TC                        | ACCACATTIC<br>TOGICITAAAG  |
|       | ACAGACTAG GCCAAAACCT TCTCTACGAT AACTSCCCC ACGCGCTAGA TGTCTGACCCCC CGCTTTTCGA AGAGATGCGT TTCACACCAG TGCGCGATCT | CHTTRACCHO CHCCCHOTHS ACCARTCOCH CHCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC | GCTCCAGTGA |            | ACACAMGCTC<br>TGTGTTCGAG | CTCTTTOAGC               |                       |            | GACCGGGGTT | CACCCTRACGT COCAACCAGG                     | ANAACATGTA               | COCCUTTINA               |  | TGGGGCCTCC<br>ACCCCGGAGG   | CCTCTAGITCT              | TTGCACTCGC<br>AACGTGAGCG | $\tilde{\mathbf{c}}$                               | TRATTROCCO GACAGACCO CACATACA CAGAACTTO CONCORCIT GAGAATTTAC   |
|       | AACTCCCCCC<br>TTGACACCCCG   | ACCAPTORCA<br>TUGINORGEGIT  | GCCCATCG   | COGCOGIACC | AACAAAGAGG               | AACATOCTAC               | TCTTCCCCCG            |            | GGAAACGGTT |  | TCTCACTTGA<br>ACAGTGAACT | TTGCCGTCTG               | AAACTCAGGC<br>TTTGAGTCCG                       | AAGTCGCAGT   | COCTUTAGE                | AGGCTTTGAG<br>TCCGAAACTC | GATICTISCTTA AAAGCCACCT<br>CTAGACGAAT TITICGGTISGA | CCTCCCTTT  |
| •     | TCTCTACGOC<br>AGAGATGCGG  | CAGGCACACA  | AACAACAGCT | TIGITICIA  | TTTCCAGGCT<br>AAAGGTCCGA | CGCACTCAAA               | CGCCATTGCT            | TTTCTCCACG | AAAGAGGTGC | ANCAGTCCCC AGGTACAGCC TTOTCAGGG TCCATCTGGG | CACTICITIT               | ACCCCCACCC               | TGCTCCACTT<br>ACTAGGTGAA<br>EcoRV              | CCATATCTTG   | CTOCCCAGCA               | CCCCCACGGG               | GATCTGCTTA   | CAGCACCTTG   |
|       | GCCAAAACCT<br>CGGTTTTTGGA   | CTTTRACCTG  | AAGCAACATC | TTCGTTGTAG | TTACAAGCGC<br>ACTCTTCGCG | COCATGGAACC              | TOCCCCCTAG            | CTCCTCCATG | GACGACGTAC | AACAGTCCCC                                 | TTAGGAGCGC<br>AATCCTCGCG | CICATIALIT               | TOGTCTTTAG<br>ACCACAAATC                       | GGTCGGGGG  | GTGGTGCACG               | CCCANANAGG               | TAMARGEETT   | GTCGTGCACG   |
|       | ACAGACCTAG<br>TGTCTAGACC  | TGTTTGAAGT<br>ACAAACTTCA  | ATAAAGAAGC | TATTTCTTCG | TCCCCTCCTA<br>ACCCCTCCAT | CTACCCCTTT               | GAGTCACTCC            | GTGGACTATT | CACCTGATAA | CTCCATGCTC                                 | AGTOCCCAGA               | ACACTETEGG               | OTTOCGATAC<br>CAACGCTATG                       | GCGTTTAGGA   | TCAGCGCCGG<br>AGTCGCGGCC | TAGCTGCCTT               | AGCGCCTGCA<br>TCGCGGACGT                           | GACAGACGC<br>CTGTCCGGC   |
|       | GGGCGCACTC  | CTTTATGTTT  | ACCCACANC  | TOCOGNOTIO | CCATATETT                | CCOTACACTO               | GITTOAGTAC            | TCOOCCOCCT | CGGA       | Kpm<br>GOOTACCCAA<br>CCCATGGGTT            |                          | TITTATTTGT               |  | CATCACCAAC   | TOGAACACTA               | TCAACTTTOG<br>AGTTOAAACC | GTTAGGATAC<br>CAATCCTATG                           | TGATTROCCG<br>ACTAACCGGC   |
|       | TTATOTCCAT  | COGOTOCCAT  | Transcract |            | 1001101000<br>ACCACACC   |                          |                       |            | CCCCGGGTTG | CTTATTACCG<br>GRATAATGC                    |                          |                          |  | CCGACGCGTG   |                          |                          | COGRETOGOE   | GCCGGNAAAC   |
|       | 21001   | 21101   | 21201      |            | 21301                    | 21401                    | 21501                 | 21601      |            | 21701                                      | 21801                    | 21901                    | 22001  | 22101  | 22201                    | 22301                    | 22401  | 22501  |

Figure ISN

| <u>+</u> < = }  | ٤ :  | · <u>- :</u> :   | <b>t</b>   | E S }   | 8 <b>8 1</b>              | . S F                    | t i                      |                             | ្រូវ                     | ៖ ដូ       | £                             | בא                       | <b>1</b> 5               | ر<br>الم الم             | : .<br>!<br>!   |
|---|--|--|--|---|---------------------------|--------------------------|--------------------------|-----------------------------|--------------------------|------------|-------------------------------|--------------------------|--------------------------|--------------------------|---|
| ATGCTTCCGT<br>TACGAAGGCA  | CNANCIGACTICS<br>OTTENCIALIA                                     | CCAGGITCTT - I   | CSCSCAGECT<br>GEGEOTICG 1  | CCTCTTCC()'V  | CCACTTO                   | AAGAAAAAGA               | GCAGGAGCCT               | MCC-NCGCGC<br>TGCAGCGC      | ACTICACETE:              |            | GOTCCTGT                      | CAGCGCCAGT               |                          | 44<br>44                 | CCTUNTATO:  |
|   | CAGTICGAGAC (  |  | CATCAGCGCG   |   |                           | TCTTCCCGCG               |                          | ACCAACCCCC                  | CTAGTACCTC               | CAGCTCCGTG | THEOTICE GOICCIGING           |                          | CACCTATTCT               | TGCTTGCCAC<br>ACGAACGOTG | CGCTCTCATA  |
| CCHTTHCCCT CCTCACATCC ATTTCAATCA CGTGCTCCTT ATTTATCATA GCAANAGCGA GCAGTGTAGG TAAAGTTAGT GCACGAGGAA TAAATAGTAT | ATCCTTCTAG C   | CCCCCCACCA   | GGTACTIGIC (CCATGAACAG   |   |                           |                          |                          | ACGIECTECA                  | CCCACATATA               | -          | TOTCTCCTAT                    | ACGTOCTOTT<br>TECACGACAA | CTACGNACGC               | GTGCCAGAGG               | ARCGGACING CAGCTGGCCT TREGGCAGGG CGCTGTCATA TCGCCTGTTC GTCGACAGGA ACGCCGTCCC GCGACAGTAT |
| ATTTCAATCA<br>TAAAGTTAGT  |  | CAGCITGCAAC  | TTATCCACGT<br>AATAGGTGCA   | CACTITICEGE   | TTTCCCATCC<br>AAACGGTACG  | GGCGGGGCT                | CCACCAGCGC               |                             |                          |            | CHCAGTACCA<br>GAGTCATGGT      | GYGGGAGACG               | TCAGCCTTCC<br>AGTCGGAACU | CGTATTINGC<br>GCATAVACGG | CAGCTGGCCT  |
| CCACTCTACATCC   | CANCCACAAC GROGAGECTA TGGGGTRGTG GRUXTHTHY CROTHEGGGG ACCGGAGGAC |  |  | ACCGTAATIT<br>TGGCATTAAA  | GCTTACCTCC                |                          | CCACACGCGC               |                             |                          |            | ACGAGGACCG<br>TGCTCCTGGC      | CTACCTAGAT               | ATAGCGGATT               | ACTICTACCC TGAAGATGGG    | AGCGGACAAG  |
| CONTINUEST  | נארונאנאל<br>הארניאלאל   | COTCACAAAG GTCTTRITTIC TGGTGAAGGT<br>GCAGTGTTC CAGAACAACG ACCACTTCGA |  | CGGGTTCATC<br>GCCCAAGTAG  | recaetatre<br>eceteracace | CCACGATTAC               | CCGCGCCCGAC              | CCTCCGCCGC                  | CTTCCCGACT               | CGCCTCCACC | ACCENTACA ACCENTACO TCCANACAT | CCGTACCGCT               |                          | CCGCGCCTCA               | ACTOCANDAT ACCOCTATICE TRACTIGITISCEN ACCRICAGE TO ACCRICATE TORICATERS                 |
| בפבטיבטינים (   | CCCCCCCAC  | COTCACAAAG   |  | DACACCATICG GCACACTICAG CGCGTTTCATT: CTGTGCTAGC CGTGTGAGTC GCCCAAGTAG | ATTCAGCCGC<br>TAAGTCGGCG  | TCCTCCCTGT<br>AGGACTGACA | ACCAGCTACC               | معمدودددرر<br>معمدودددرر    | COCTOCTCCT               |            |                               | GGGGACGAAA               |                          |                          | TACCGTTSCCA   |
| CTCCTTCAGC (  | TCGATTTTAG   | GCCCCATCAT   |  | Pwd<br>GACACGATCG<br>CTGTGCTAGC                                       | CCAGCAGAAG                | TECTETATET               | TCCOCCGCCG<br>AGGCGGCGCC | <b>GCTTTTTTCG CGANNAACC</b> | CCACCAAAGC               | CCCTCTGAGT | AGCAGGACCC<br>TCGTCCTGGG      | AGTCGGGCGG               |                          |                          | ACTOCANGAT ACCCCTATCC TOCKGTGCCA  |
| TRICTAGACTR C   | MAGCTCGCCT TTCGAGGGA   | Pstl<br>TOCAGGAATC   |  | CTCCCACGCA  | CCCCCACTO                 | GCGCCACATC               | ANTGGCCAAA               | CGCCTCATCC                  | COCOCTCOGO               | CCTAACCGCC | OTGATTATCO<br>CACTAATAGC      | ACCAGGAACA               |                          | _                        |   |
| ATCTTGGCCT TGCTAGACTG CTGCTTTAGC GGGGGGTAGC<br>TAGAACCGGA ACGATCTGAC GARRANGTCG GGGGGAGACGG                   | GTAGACACTT /   | Pstl<br>CAGGTACGCC   | CATACOCCO CONTRACTOCCO CONTRACTOCO | CCATGCCCTT  | CCCCATACCA                | ACCATTIGIA               | TCTTTOGGCGC              | CTCGATACGC                  | OCACCOCOTO<br>COTOOCOCAO | AGAAGGACAG | CCTCCTCCTT                    | GCAGAGGCAA               | GCGCCATTAT               | ACCCCCCAAA<br>TCCCCCCAAA | THITTICCAAA   |
| 22601   | 22701  | 22801  | 22901  | 23001   | 23101                     | 23201                    | 23301                    | 23401                       | 23501                    | 23601      | 23701                         | 23801                    | 23901                    | 24001                    | 24101   |

Figure 150

## PMRKAdSgag MER682

| COCCOCANO TOTALONGERA  GOGOGICTEC AGTACTECTO TCACTACTECT  ATGCCGTCA ACCGCTOCTO CACCTACTCC  Byll  COCCAGGCCT GAAGATCT  COCCAGGCCT GAAGATCT  COCTCAAGG CTTCAAGG GTCACTCG  COCTCAAGG CTTCAAGG GTCACTCG  COCTCAAGG CTTCAAGG GTCACTCC  COCTCAAGG CTTCAAGG GTCACTCCT  COCTCAAGG CTCCCCGC GCGACTACC  GOAGGTCCC GCTCCCGCG GCCTCATCC  GCACCTCAAGG CTCCCCAAGG AGCTCCTATC  GAAGATCATT TCCCCGAAGG CCTGCCAAA  CTCCTCAAGG CTCCCAAGG AGCTCCTATAC  GAACATCATT TCCCCGAACG CCTGCTTAAA  CTCTCTAAGG CCTCCCAAGG CCTGCTTAAA  CCTCCTTAAAA AGGGGCTTCC  GAACATCATT TCCCCGAACG CCTGCTTAAA  CTTCTAAAAA AGGGGCTTCC  GAACATCATTT  CACCTTAAAA  CCTCCTTAAAA   CCTCCTTAAAAA  CCTCCTTAAAAA  CCTCCTTAAAAA  CCTCCTTAAAAA  CCTCCTTAAAAA  CCTCCTTAAAAAA  CCTCCTTAAAAAA  CCTCCTTAAAAA  CCTCCTTAAAAAAAA |
|--|
| AGG ANTETTGCCC  JEGIN  JENA  JENA  GANGGENACT  GATC GATCGTTGA  THG GANGGGNCA  THG GANGGGNCA  THG GANGGGNCA  THG GANGGGNCA  THG GCCCCCCTG  THG GCCCCCCTGA  THG GCCCCCCTGA  THG GCCCCCCCTGA  THG GCCCCCCCTGA  THG GCCCCCCCTGA  THG GCCCCCCCCTGC  THC CANAGGGGCGA  THC CANAGGGGCGA  THC CANAGGGGCGA  THC CANAGGGGCGA  THC CANAGGGCGCA  THC CANAGGGGCGA  THC CANAGGGGCGA  THC CANAGGGCGCA  THC CANAGGGCGCA  THC CANAGGGCGCA  THC CANAGGGCGCCCCCCCCCCCCCCCCCCCCCCCCCCCCC  |

Figure 15P

pMRKAd5gag MBR682

| 25701 | GGGCCTTGC   | Treccagant  | GOGCCCTTGC TPCCCABBAT CCCACCTAA AAGAACCTTGC NETTETCTCCCCCCCCCCCCACGACC CCCCCCCCCAA TTCTTCCACC TTCACCCCCCCCC | AAGAAGCTGG 1                   | M.H.THRETYGEC<br>PYCANYAGEGGG | Treccadant occaceras arganetric altertering gettrectare grichard arganetric altertering francering checker construit tretregace treateness construit trecretices | TACGAGGAGG A        | AATACTGGGA C                                   | CAGTCAGGCA     | GAGGAGGTTFF CTCCCA . |
|-------|-------------|-------------|---|--------------------------------|-------------------------------|--|---------------------|--|----------------|----------------------|
| 25801 | TUCACGAGGA  | CCACCACCAC  | ATGATGGAAG  | ACTROGRAMA CITTAGACIAN         |                               |  | _                   |  |                | CACCCTCG/            |
|       | ACCTGCTCCT  | cerecreero  | TACTACCTTC  |                                |                               |  |                     |  | CTTTGTGCA      | Ulter Michigan       |
| 25901 |             | Tegecogece  |   |                                | TCCACCATOG                    | CTACAACCTC C   | COUNTRICATE C       |  | GTCACGGCA      | ACCOCTOR             |
| 10030 | GCGTAAGGGG  | ACCOUNTS    | TESTABLISHED  | COUNTRACTOR OF THE PROPERTY OF |                               |  |                     |  | CCAMBOCTAC     | COCICATOO!           |
| 10007 | TTOOCATCTA  | CCCTCTCCTO  |   |                                |                               |  |                     |  | GGTTCCGATG     | OCGAGTACC!           |
| 26101 | OCOCOCACAA  | GAACGCCATA  | GTTGCTTGCT  | TOCANGACTO                     | ACCCCCGTTG                    | ATCTCCTTCG (TAGACGAAGC   | CCCCCCCCTT GOCCCCAA | TCTTCTCTAC AGAMANGATE                          | CATCACGGCG     | ACCTATANGG".         |
| 26201 | CCOTAACATC  | CTGCATTACT  |   | CTACAGCCCA                     |                               | _  |                     |  |                | מאכמאככסמא           |
|       | OCCATTICTAG | GACGTANTGA  | TCCCAGTAGA  | GATOTOGGGT                     | ATGACGTGGC                    | -  |                     |  |                | ברנירופינרו          |
| 26301 | TAGCAAGACT  |             | CCAAGAAATC  | CACAGCGGCG                     | GCARCAGCAG                    | GAGGAGGAGG C   | GCTCCCTCTG C        | COCOCCANCO                                     | TTOGGCATAG     | CTGGGCGCTC           |
| 26401 | ATCOTICION  | GACTURITION | CACTESTAS   | CCTATATTE                      |                               |  |                     |  | CAGGICTOTO     | CCATCCCTCA           |
| 10507 | GAATCTTTGT  |             | GTOROACATA  | CGATATAAAG                     |                               | _  | GTTCTCGACT          | -  | GTCCAGAGAC     | GCTAGGGAGT           |
| 26501 | CCCGCAGCTG  |             | NAAAGCGAAG  | ATCAGCTTCG                     | GCGCACGCTG                    |  |                     | CACTAAATAC                                     | TOCOCOCTOR     | CTCTTAAGGA           |
|       | GOOCGICGAC  |             | <b>trincectic</b>   | TAGTCGAAGC                     | CGCGTGCGAC                    | CTICIGCGCC   |                     | GICATITATO                                     | ACCCCCCACT     | GAUNATTICC           |
| 26601 | CTAGTITICOC | OCCUPTICAC  | AAATTTAAGC  | CCGANAACTA                     | CGTCATCTCC                    |  |                     | GCACCTGTTG                                     | TCAGCGCCAT     | TATGACCIAG           |
|       | GATCAAAGCG  |             | TITAMATICG  | CGCTTTTTCAT                    | ACAGTABAGG TCGCCCGTGT         |  |                     | CCTCGACAAC                                     | AGREGEGOTA     | AFACICEITC           |
| 26701 | GANATICCCA  | COCCCTACAT  | GPGGAGTTAC  | CACCCACAM                      | TGGGACTTGC                    |  | GCCCAAGACT          | ACTCAACCCG                                     | AATAAACTAC     | ATGAGCGCGG           |
|       | CTITAAGGGT  | OCCOGGATGTA | CACCTCAATG  | GICCOTGITT                     | ACCCTGAACG                    | ACCENTAACO CCGACCTCGA  | COGGITTCTGA         | TCACTTGGGC                                     | Tratticate     | TACTURGEU            |
|       |             | Ecofiv      |   |                                | Ψ ¥                           | Ecofil   |                     |  | ,              |                      |
| 26901 | GACCCCACAT  | -           | GTCAACGGAA  | TACCCCCCA                      | CCCHANCCGA                    |  | MACAGGGGGG          |  | ACACCTCOTA     | ATAACCTTAA           |
|       | CTOCOCOTOTA |             |   | ATGCGCGGGT                     | GOCTFTOGCT                    | CACTITICS TANGAGGACC   |                     |  | TOTOGRACIA     | TATTCGAATT           |
| 26901 | TCCCCGTAGT  | TOCCCCCTG   | CCCTGGTGTA  | CCAGGMAAGT                     | CCCCCTCCCA                    |  |                     |  | CCGAAGTICA     | GATCIACTIAN          |
|       | AGGGGCATCA  | ACCOMICTOR  | GOGACCACAT  | GGTCCTTTCA                     | CARCCOACCO                    | GGTGACACCA   | TCAAGGGTCT          |  | GGCTTCAAGT     | CTACICIATIN          |
| 27001 | TCAGGGGCGC  | AGCTTGCGGG  | COCCUTICGE  | CACAGOGING                     | GOTCGCCGG                     |  | ACTICACCTGA         |  | GCCAGGTATT     | CAGCTCAACT           |
|       | AGTCCCCCCC  |             | GCCGNANGCA  | GRETCCCACG                     | CCAGCGGGC                     | CCTCCCATAT   | TGAGTGGACT          |  | CCCTCCATAA     |                      |
| 27101 | ACGAGICOGE  | GAGCTCCTCG  | CPTOCITCTCC   | GTCCGGACGG                     | GACATTTCAG                    |  | ככנוטככטכעכ         |  | CCTCGTCAGG     |                      |
|       | TRECTCAGCCA |             | GAACCAGAGG  | CAGGCCTGCC                     | CTGTAAAGTC                    | TAGCCIACCAC  | GGCCGGCGAG          | AAGTAAGTGC                                     | GGAGCAGTUC     | GTTAGGATTG           |
|       | Pell        |             |   |                                |                               | ÷  |                     |  |                |                      |
| 27201 | TCTGCAGACC  | regreeners  | ACCCCCCCCC  | TOGAGOCATT                     | GGAACTCTGC                    | TETOCAGACE TESTECTETS AGGEGRETE TOSAGACATT GRAACTETEC AATTTATTGA GAAGTTTGTG  | CCACTTGTG           | CCATCOGICT ACTITION COLOCIC COLOCIC CANADACTOR | ACTITION       | CHICAGOS             |
|       | ACACCTICAG  | ACCAGGAGAC  | TCCCCCCAG   | ACCTCCGTAA                     | CCTTCAGACT                    | ARCARDAGA TOGGOGGAG ACCTOCOTAA CCTTYAAAAA TAAATAACT CCTCAAACAC GGTAACCAGA TAAAATTAAG WAXAAACAC   | CCTCAMCAC           | GGIMELAUA                                      | Service I town |                      |

Figure 1501

### DMRKAd5qag MER68

| 27101            | CCTCCGGCC     | ACTATCCOOA | TEAATHTAIT   | CCTAACTTAG             | ACCCCCTANA        | מפענבנטטכט (  | GACGGCTACG     |                    | NACTICGNGAG  | BCAGACCAAC    |
|------------------|---------------|------------|--------------|------------------------|-------------------|---------------|----------------|--------------------|--------------|---------------|
| 1                |               |            | AGTTANATAN   |                        | TREVERSE          | CETTGAGETGE   | כווטטירהאדמכ י | THACTTACA          | TTCACCTCTC   | CONCINCIATION |
| 27401            | •             |            | CACTGTCGCC   |                        |                   |               |                |                    |              | ATATCGAGGG    |
|                  | ACCCCCCACTT   | TGTGGACCAG | GTGACAGCGG   | COOTOTTCAC             | מאארדוממכ         |               |                | _                  |              | TATABLICA     |
| 27501            | CCCGGCGCAC    | 995975595  | TTACCGCCCA   | GOGAGAGCTT             | GCCCCTAGCC        | TCATTCGGGA    |                |                    |              | GGACACKO:A    |
|                  | 0             | CCGCAGGCCG |              | CCCTCTCGAA             | CCCCCATCGG        | <b>ACCCCT</b> | CAAATGGGTC     | <b>OCCOGOOGNCO</b> | ATCAACTCGC   | ccroncec      |
|                  |               |            |              |                        |                   | rgii!         |                |                    |              |               |
| 27601            | cccroromc     | TCACTGIGAT | THICHACTET   | CCTAACCCTG             | GATTACATCA        |               |                |                    | TAATAAATAC   | AGNANTTAA     |
| 1                | GCGACACAAG    | ACTOACACTA | AACGITICACA  | GGA-FTGGGAC            | CTAATGTAGT        | TCTAGAAACA    | ACCCTAGAGA     | CACGACTCAT         | ATTATTATO    | TCTTTAAT      |
| 27701            | ATATACTOO     | OCTCCTATCG | CCATCCTGTA   | AACTICCACCO            | TCTTCACCCG        |               |                | CCTTACCTOG         | TACTITITAAC  | ATCTCTC: .    |
| <br> -<br> -<br> | TATATOACCC    | CGAGGATAGC | _            | TRECOGNAGE             | AGANGTORGC        | GOOTTCOTTT    | GGTTCCCCTT     | CONNTCANCO         | ATGAMATTO    | TMEMGAGGEN    |
| 27801            | CHUICATTEA    | CAACAGITIC | AACCCAGACG   | GAGTGAGTCT             | ACGAGAGAAC        | CTCTCCGAGC    |                | CATCAGAMAA         | AACACCACCC   | TCCTTACC":1   |
| )<br>}<br>•      | GACACTANAT    | GTTGTCAAAG | TYCCCTCTCC   | CTCACTCAGA             | TECTCTCTTG        | GAGAGGCTCG    | AGTCGATGAG     | Gragactifit        | rrerocrose   | AGGNATGGAC    |
| 27901            | CCOGGRACOT    | ACGAGTGCGF | CACCOCCCC    | TREACCACAC             | CTACCGCCTG        | ACCOTANACC    |                | COGACAGACC         | TCAATAACTC   | TGTTTACCAG    |
| •                | GCCCTTOCA     | TOCTCACGCA | GTGGCCGGCG   | ACGTOCTOTO             | GATCCCCGAC        | TOCCATTIFG    | TCTGAAAAAG     | GCCTGTCTOG         | AGITTATTICAC | ACAMATKAGIT   |
| 28001            | AACAGGAGGT    | GAGCITAGAA | AACCCTTAGG   | GTATTAGGCC             | ANARGEGEAG        |               | GTTTATGAAC     | AATTCAAGCA         | ACTICITACOGG | CTATTCTAAT    |
|                  | TTGTCCTCCA    | CTCGAATCTT | TTGGGAATCC   | CATAATCCGG             | TTTCCGCGTC        | CATCACACCC    | CANATACTTO     | TTAAGTTCGT         | TGAGATGCCC   | GATAAGAT"     |
|                  |               | Kbsl       |              |                        |                   |               |                |                    |              |               |
| 28101            | TCAGGITTICS   | CTAGAATCGO | COTTICGGGTT  | ATTICTOTO              | TIGHTOTICE        | CHITAINICHT   | ATACTAACOC     | TICICICACCE        | AAGOCTCGCC   | OCCIDENT:     |
|                  | AGTECCAAAGA   | GATCTTAGCC | CCACCCCAA    | TAAGAGACAG             | AACACTANGA        | GANATAAGAA    | TATGATTGCG     | AAGAGACGGA         | TTCCCAGCGG   | CCCACCACAC    |
| 10000            | Section 1     |            |              |                        | TCGCCACCCA        | AGATGATTAG    | GEACATAATC     | CTAGGTTTAC         | TCACCCTTGC   | OTCAGCCAC     |
| 10907            | ACCITOTANAC   |            |              |                        | AGCGGTGGGT        | TCTACTAATC    | CATGTATTAG     | GATCCAAATG         | ACTOGGAACG   | CAGTCGGGTV1   |
|                  | Kpr#          |            |              |                        | -                 | -             |                |                    |              |               |
| 28301            | GOTACCACCC    | AAAAGGTGGA | TITITANGGAG  | CCAGCCTGTA             |                   |               |                | GCACCACTCT         | TATAMATGE    | ACCACAGO      |
|                  | CCATOGTOGG    | -          | - AAAATTCCTC | GGTCCCACAT             | TACAATGTAA        | GCGTCGACTT    |                | CCTCGTGAGA         | ATATITIACO   | regrenting.   |
| 10786            | ATTORAGA      | _          | CACMANACA    | ANATTOCCAA             | GTATCACTUST       | TATGCTATT     | CHICAGCCAGG    | TCACACTACA         | GACTATAATO   | TTACAGITT     |
|                  | TACTTITICGA   |            |              | TTTAACCGTT             | CATACGACAA        | NTACGATAAA    | ccorcoorcc     | ACTGTGATGT         | CTCATATTAC   | AATGTCAAAA    |
|                  |               |            | Hsil         | Bs111071               |                   |               |                |                    |              |               |
| 28501            | CLACACACATAAA | ACTICATARA | _            | CHITTATOTA TACTITITICA | TITITATOWN        | TGTCCGACAT    |                |                    | AGTATAGGTT   | GTGGCCCCCA    |
| 4                | COTCCCATTT    |            |              | ATGAMAGGT              | MANTACTITE        | ACACGCTGTA    | ATGGTACATG     | TACTCONTING        | TCATATTCAA   | CACCGIRAGUT   |
| 28601            | CAAAATTGTG    | -          | TOCACTITIC   | TRCTCCACTG             | CTATGCTAAT        |               |                | GTACCCTACT         | CTATATTAAA   | TACAMAGCA     |
| !<br>!<br>!      | GTTTTAACAC    | _          |              | ACGACGTGAC             | GATACGATTA        | ATCTCACGAG    | CCANACCAGA     | CATGGGATGA         | GATATAATTT   | ATCLLIC       |
| 28701            | GACISCAGCTT   |            |              | CTTAATTTAC             | TANCITACAA        |               |                | GCTTTACTCO         | CTCCTTCCAA   | AACAAATT      |
| :<br>:<br>!      | CTOCCTCGAA    |            |              | GAATTAAATO             | ATTCAATGTT        |               |                |                    | GACGMACGIT   | TICH TANK     |
| 78801            | AAAGTTFAGC    | ATTATATTA  | GAATAGGATT   | TAMCCCCCC              | <b>GCTCATTTCC</b> |               |                |                    |              | ATATISCICA    |
| ;                | TITTICAATCO   |            |              | ATTTGGGGG              | CCAGTANAGG        | ACGAGTTATG    | GTAAGGGGAC     | TTGTTANCTG         | AGNTACACCC   | TATACCACOT    |

# PMRKAd5qag MEH682

| ,      |             | STATE TO THE STATE OF THE STATE |                       | ATCTONICATE OF                                 |  | CCAGGACCTG T  | TCCCGCCGCAT 1                    | THURSTICCAGE (        | CCAACTACAG CGACCCACCC | KINCCARIC                  |
|--------|-------------|--|-----------------------|--|--|---|----------------------------------|-----------------------|-----------------------|----------------------------|
| 10697  | GCGC INCHES | ACCOMPANY OF THE STATE OF THE S | CLTSANCIGALT          | TACAGTEGTA GACTGAAAGE                          |  |   |                                  | AACAAGGTCA (          | GGTTGATGTC (          | SCHLADOTH ST               |
| 29001  | TAACAGAGAT  | TAACAGAGAT GACCAACACA ACCAACGCGG   | ACCAACGCGG            | כבפבבפינדעב                                    |  |   | ATACACCCCA I                     | AGTITICHOCC 1         | AAACAGTTAT            | ACTYGGAT'AA<br>TGACCCTA'IT |
|        | ATTOTOTOTA  | CTGGTGTGT  |                       | TEGITALIST, CALINDONIA OCCUMINA                |  |   |                                  | _                     | ACGCGCCCG             | ACCACCCATC                 |
| 29101  | CHICOCCATO  | TOGTOGETEET  | CCATAGGGGT            | ATACAAACAT                                     |  | _   |                                  | _                     | -                     | TOGTVICKITAG               |
| 29201  | TATAGECCA   | TCATTOTGCT   |                       | ACACCCAAAC AATTIATTIAA TEETATAGATT GAACGCAATTG | TECATAGATT   | -   |                                  |                       |                       | TTANATCIACIA               |
| ,<br>, | ATATCAGGGT  | AGTAACACGA   |                       | TTACTACCIT                                     | TINCTANCIT AGGINICIAN CCTGCCTGAC                       | -   | TTTGTGTACA                       | ACARAGAGA             | ATGTCATACT            | AATTTACTO                  |
|        |             | Mhol westernooned to the contract of the contr |                       |  | i di i di dalah da da da da da da da da da da da da da |   | ATTOGCTOCG                       | GTFFCTCACA            | TCGAAGTAGA            | CTGCATTC: A                |
| 10562  | GTACTAAGGA  | CATCATICCY CCACIFILIA  | ATAATGACTG            | ATANTGACTE GGAACAACGC GAANANACAC GCACGAGGTG    | CAAAAAAACAC  |   |                                  | CAAAGAGTGT            | AGCTICATET            | GACGTAAG .T.               |
|        |             |  |                       |  | PsI  | _   |                                  |                       |                       |                            |
| 29401  | OCCUTICACAG | <b>1CTATTIGCT</b>  | TTACGGATTT            | GICACCCTCA                                     | CGCTCATCTG   |   | ACTOTICATE TECCEPITAT CEASISCATT | TCCCCTTTAT            |                       | CHSACCEAGA                 |
|        | COGAROTOTIC | AGATAAACGA   | AATGCCTAAA            | CACTGGGAGT                                     | GCGAGTAGAC GICGAMOTAG                                  |   | Ecoff                            |                       |                       |                            |
|        |             |  |                       | TOROGRAPHO CASTATORICA                         |  | ATAGCTTCAGC   | TYCTTAGAAT TCTTTAATTA            |                       | TOMATITIAC            | TOTOACTT!                  |
| 29501  | CACACCCCTT  | ACCATATICAL ACCAMBAGA  |                       |  | CCTGTCCTGA   |   | ANGANTETTA AGAMATTANT            |                       | ACTITABATO            | ACACTONANA                 |
| 10505  | CHCACGCGA   |  |                       | ATCHEROTER DETRECCEGA CETECAAGCE               | CCTCCAAGCC   |   | ATATCATGCA                       | GATTCACTCG            | TATATGGAAT            | ATTCCAAGIT                 |
| 70067  | GACGACTAAT  |  |                       | ACANGGGGGT                                     | ACANGOGGET GGAGGETICEG                                 |   | TATAGTACGT                       | CTAAGTGAGC            | ATATACCITIA           | TAAGGTTCAA                 |
|        |             |  |                       |  |  |   | Pstl                             |                       |                       |                            |
| 29701  | CCTACAATGA  | AAAAAGCGAT   |                       | CITICGGAAG CCTGGTTATA TGCAATCATC               | TCCAATCATC   | TCTGTTATEG '  |                                  | TACCATCITA            |                       | TATATCCC                   |
|        | CGATGITACT  |  | GAAAGGCTTC            | GCACCAATAT                                     | ACCITAGTAG   | ACCTTAGTAG AGACAATACC ACAAGACGTC                                    |                                  | ATGGTAGAAT            |                       | ATATMAKIAT                 |
| 29801  | CCTTGACATT  | GOCTOGRACG   | CANTAGATGC            | CATGAACCAC                                     |  | CCANCTITICE CCGCGCCCGC TATGCTTCCA                                   |                                  | CTCCAACAAG            | TIGHTGCCGG            | COCCTTTCTC                 |
|        | GGAACTGTAA  |  | GTTATCTACG            | GTACTTGGTG                                     | CONTRANCO  | CONTRADAGO GOCGCGGCG ATACGAAGGT                                     |                                  | GACGITIOTIC           | AACAACGGCC GCCGAAACA  | GCCCBAARCA I               |
|        |             | •  |                       |  |  |   |                                  |                       | BgH                   | •                          |
| 10000  |             |  | N. P. STATE PROPERTY. |  | AAATCAGCTA   | ACCICIONIN AANTCARCIA CITTAANCIA ACARAARRAGA ANGACIBACA             | ACAGGAGGAG.                      | ATGACTGACA            |                       | AGAAATGGAC                 |
| T0667  | CONCERNIO   | -  |                       |  | TITAGIKIGAT  | TITAGIN:GAT GAAATTAGAT TOTCCTCCTC TACTGACTGT                        | TOTOCTOCTO                       | TACTGACTGT            | GGGNTCTAGA            | TCTTTACCTG                 |
| 10001  | CESANTIALIA |  | _                     | AGACGCAGGG                                     | CADCODCCOA GCNACAGCGC                                  |   |                                  |                       | CATOOTTAAC            | TICCACCAGE                 |
|        | CCTTAATAAT  |  | GCACGATCTT            | TCTGCGTCCC                                     | GTCGCCGGCT   | GREGGEGET COTTGTEGES  | TACTFAGTTC                       | TCGAGGITICT           | GTACCAATIG            | AALTHUGILA                 |
| 10101  | CEAAAAAGGGG |  | CTCGTAAAGC            | _  | CACCTACGAC   |   | CCGGACACCG                       | CCTTNGCTAC            | AAGTTGCCAA            | CCANGCOTY                  |
|        | CONTINUENCE |  | GACCATTICG            |  | PECCEPTIFICA GIUGATICCIU TCATTATUCET                   | TCATTATEST  | GCCTGTGGC                        | GGANTCGATO            | TECACGGT              | GET W.C. M.:               |
| 10201  | CANATTROCTO | _  | GAGANAAAGCC           |  | CATTACCATA ACTCAGCACT CRATAGANAC                       |   | CCAAGGCTGC                       | AFFICACTCAC           | CTTCTCAAGG            | ACCTGARGAT                 |
|        | CTTTANCCAC  | _  |                       |  | TCACTCGTGA   |   | GCTTCCGACG                       | TAAGTGAGTG            | GAACAGITICC           | TECACICETA                 |
|        |             |  |                       | Bylls<br>seeming                               | ł  |   |                                  |                       |                       |                            |
| 30301  | CTCTGCACCC  |  | : ccrorecoor          | CTCAAAGATC                                     | TTATTCCCTT   | CHENANGATE: TTATTCCCTT TAACTAATAA AAAAAATAA TAAAGCATCA CITACTTAAA   | ANAMANTAA                        | TRANGCATCA            | CTTACTTAAA            | RICAGII (A.)C              |
|        | GAGACGTGGG  |  | GGACACGCCA            | GAGTITICTAG                                    | MATANGGGMA   | ANTANTICTE GENEACECEA GAGTITETAG ANTANGORIAN ATTENTIATE TITITITIATE | ተተተተተተ                           | ATTICCTAGT GAATGAATTT | GANIGAAT              | TAUTICAMICO                |

## PMRKAd5gag MER682

| 30401 | ANATHETET    | CCAGTTTATT                              | CACCACCACC  | rectronect            | CCTCCCAGGT            | CHARTATION            | AGCTTCCTCC   | TOCCTOCAAA            | CTITICACAC         | ANTCTARATE        |
|-------|--------------|---|-------------|-----------------------|-----------------------|-----------------------|--------------|-----------------------|--------------------|-------------------|
|       | TTTANAGACA   | GGTCAAATAA                              | Greencenee  | ACCAACCCCA            | CCACACATECA           | GACCATAACG            | Trgmagaag    | ACCGACOTITE           | GAAAGAGGTG         | TTACATTTAC        |
| 30501 | CAATOTCAGT   | Trechecter                              | TCCTGTCCAT  | CCGCACCCAC            | TATCTTCATG            | Therrecash            | TGAAGCGCGC   | AACIACCGTCT           | GAAGATACCT         | TCANCCCCCT        |
|       | CTTACACTICA  | AAGGAGGACA                              | AGGACAGGTA  | GGCGTGGTG             | ATAGAAGTAC            | AACAACGTCT            | ACTITORICACG | TTCTGGCAGA            | CTTCTATOGA         | ACTTGGGG A        |
| 30601 | GTATCCATAT   | GACACOGAAA                              | CCGGTCCTCC  | AACTGTGCCT            | TTTCTTACTC            | CICCCTITGE            | ATCITCCCAAT  | <b>GGGTTTCAAG</b>     | AGAOTECECE         | TGGGGTAC1:        |
|       | CATAGGTATA   | CTOTOCCIPIT                             | CCCAGGAGG   | -                     | TTGACACGGA AAAGAATGAG | GAGGGAAACA            | TAGGGGGTTA   | CCCAAAGTTC            | TCTCAGGGG          | ACCCCATGAG        |
|       |              |   |             | ils:                  | it i                  |                       |              |                       |                    |                   |
| 30701 | Termococc    | TATCCGAACC                              | TCTAGITACC  | -                     | TCCAATGCCA TGCTTGCGCT | CAAAATGGGC            | AACTOCOLOTOR | CTCTGGACGA            | GGCCGGCAAC         | CTTACCTCC(:       |
|       | AGAAACGCGG   | ATAOCCTTOO                              | AGATCAATGG  |                       | AGGTTACCGT ACGNACGCGA | GTTTTACCCG TTGCCGGAGA | TTGCCGGAGA   | GAGACCTUCT            | ccooccorro         | GANTGGAGE 1       |
| 30801 | AAAATOTAAC   | CACTOTORGE                              | CCACCTCTCA  |                       | AAAAAACCAA GTCAAACATA | AACCTGGAAA            | TATCTOCACC   | CCTCACAGTT            | ACCTCAGAAG         | CCCTAACTEST       |
|       | TITITACATIO  | GTOACACTCO                              | COTOCACACT  | TTTTTTT               | CAGTITICIA'F          | TTCROACCTIT           | ATAGACGTOG   | CGAGTGTCAA            | TOGAGICTIC         | GOODTIGACA        |
| 30901 | 99CT9CC9C    | GCACCTCTAA                              | TGGTCGCGGG  | CAACACACTC            | ACCATGGAAT            | כאכאניסכבככ           | GCTAACCGTG   | CACGACTCCA            | AACTTAGCAT         | TOCCACCCAA        |
|       | CCGACGGCGG   | CGTGGAGATT                              | ACCAGCGCCC  | GTTGTGTGAG            | TIGGTACCITTA          | GTGTCCGGGG            | CCATTCCCAC   | GTGCTGAGGT            | TTGAATCGTA         | ACGGTGGGT I'      |
| 31001 | OGACCCCTCA   | CAGTGTCAGA                              | AGGNANGCTA  | GCCCTGCAAA            | CATCAGGCCC            | CCTCACCACC            | ACCGATAGCA   | GTACCCTTAC            | TATCACTOCC         | TCACCCCCTT        |
|       | CCTOCOCAGE   | GICACAGICE                              | TECTTICGAT  | COCCACCTTT            | GTACTCCOCG            | CCACTCCTCC            | TOGCTATOGT   | CATOGGANTO            | ATACTICACOO        | AGTGGGGGA1        |
| 31101 | TAACTACTOC   | CACTGGTAGC                              | TTGGGCATTG  | ACTTGAAAGA            | OCCCATITIAT           | ACACANANTO            | GAMMCTAGG    | ACTARAGIAC            | OCCOUNT            | TOCATGEM :        |
|       | ATTGATGACG   | GTGACCATCG                              | AACCCGTAAC  | TOANCIFICE            | COCGINAATA            | TGTGTTTTAC            | CITITIGATEC  | TOATTICATO            | CCCCGAGGAA         | ACCITACATTIVE     |
| 31201 | AGACGACCTA   | AACACTITIGA                             | CCGTAGCAAC  | TOGITCCAGGT           | CTCACTATTA            | ATAATACTTC            | CTTGCAAACT   | AAAGTTACTO            | GAGCCTTGGG         | TITIOATICA        |
|       | TCTGCTGGAT   | TTOTOAAACT                              | OCCATEGING  | ACCAGGICCA            | CACTGATAAT            | TATTATGAAG            | GAACCITITGA  | TITICAATUAC           | CTCGGAACCC         | ANANCTAN .        |
| 31301 | CAAGGCAATA   | TOCANCITAN                              | TĠTAGCAGGA  | OCACTAAGGA            | TRATTCTCA             | MACAGACCC             | CTTATACTTO   | ATCITACTEA            | <b>TCCONTINGAT</b> | <b>GCTCAAAACC</b> |
|       | GITCCOTTAT   | ACGITICAATT                             | ACATCGTCCT  | CCTGATTCCT            | AACTAAGAGT            | THOTOTOCG             | GAATATGAAC   | TACANTCAAT            | AGGCAAACTA         | CONCTITION        |
| 31401 | AACTAAATCT   | ANGACTADGA                              | CAGOGCCCTC  | TITITIATAAA           | CTCAGCCCAC            | AACTTGGATA            | TTAACTACAA   | CANAGOCCITY           |                    | CAGCTTCAAA        |
|       | TTGATTTAGA   | TICTOATICCT                             | GTCCCGGGAG  | ANAANTATIT            | GAGTCGGGTG            | TAGAACCTAT            | AATTGATGTT   | GIFFICCOGAA           | ATGAACAAAT         | GTCCAAGTT F       |
|       |              | Handia                                  |             |                       |                       |                       |              |                       |                    |                   |
| 31501 | CAATTCCAAA   | AACCTTGAGG                              | TTAACCTAAG  | CACTGCCAAG            | GOGTTCATGT            | TTGACGCTAC            | AGCCATAGCC   | ATTAATGCAG            | GAGATGGGCT         | TOMPTTON          |
|       | GTTAAGGTTT   | TTCOAACTCC                              | AATTCGATTC  | GTGACGCTTC            | CCCAACTACA            | AACTGCGATG            | TOGGTATOGG   | TAATTACGTC            | CTCTACCCGA         | ACTITANACCA       |
| 31601 | TCACCTAATO   | CACCAMACAC                              | AAATCCCCTC  | MAAACMMAA             | TTGGCCATGG            | CCTAGAATTF            | GATTCAAACA   | AGGCTATGGT            | TECTAMACTA         | GOLACTICAT        |
|       | ACTOGRATIFAC | orcorrroro                              | TTTAGGGGAG  | TITICITITE            | AACCGGTACC            | GGATCTTAAA            | CTAACTTTGT   | TCCGATACCA            | ACCATTTCAT         | CCTTGACCCK        |
| 31701 | TTAGTTTTOATT | CAGCACAGGT                              | GCCATTACAG  | TAGGAAACAA            | ANATAATGAT            | AAGCTAACTT            | TYTTYCACCAC  | ACCAGCTCCA            | TETECTANCT         | GTACACTNAA        |
|       | AATCAAAACT   | GICOTOTICCA                             | CGGTAATGTC  | Aircomment            | TITATITACTA           | TICGATTGA             | ACACCTRIGAG  | TOCTCCAGGT            | ACACCATTICA        | CATCTGATTT        |
| 31801 | TOCAGAGANA   | GATGCTAAAC                              | TCACTITIOGE | CTTAACAAA             | TGTCXCAGTC            | MANTACTTGC            | TACAGITITCA  | GTTTTGGCTG            | TTAAAGGCAG         | THIGGENECA        |
|       | Acotetern    | CTACGATITIG                             | AGTGAAACCA  | GANTIGITITE           | ACACCGTCAG            | TTTATGAACG            | ATCTCAAAGT   | CAAAACCGAC            | AATTICCGIC         | AAACCGAGGT        |
| 31901 | ATATETOGAA   | CAGTTCAAAG                              | TOCTCATCTT  | ATTATAGAT             | TTCACGAAAA            | TREAGTGCTA            | CTANACAATT   | CCTTCCTGGA            | CCCAGAATAT         | Tr. JAACTITA      |
|       | TATAGACCTT   | PATAGACCTT GTCAAGTTTC                   | ACGAGTAGAA  | TANTATTCTA            | ANCTECTTIT            | ACCTCACGAT            | CATTICITAA   | GGAAGGACCT            | GCOTCTTATA         | ACCTITIONAL       |
|       | 9            | Jan Jan Jan Jan Jan Jan Jan Jan Jan Jan |             |                       |                       |                       |              | -                     |                    |                   |
| 32001 | GAMATOGAGA   | GARATGONGA TCTTACTGAA                   |             | OCCACACCT ATACANACCC  | TGTTGGATTT            | ATGCCTANCC            | TATCAGCTTA   | TCCAAAATCT            | CACOOTIVAAA        | CTGCCAAAAG        |
| :     | CITTACCTCT   | CITTACCTCT AGAATGACTT                   | CCGTGTCGGA  | TATOTITICG ACAACCTAAA | ACAACCTAAA            | TACGGATTOG            | ATAGTCGAAT   | AGGITTINGA GIGCCAITIT | GTGCCATTTT         | GACCOCTTITIC      |

Figure 15T

# PMRKAd5gag MER682

| 10161 | PAACATHETE   | ACTIVABILITY | ACTTANACCE   | ACACAAAAACT           | AAACCTCTAA                 | CACTAACCAT ' | TACACTANAC                | GCTACACAGG  | NNCAGGINGA        | CACAACTOUA    |
|-------|--------------|--------------|--------------|-----------------------|----------------------------|--------------|---------------------------|-------------|-------------------|---------------|
|       | ATTICTAACAG  |              | TGAATTTGCC   | TOTOTAL               |                            |              | ATGREATITG                | CCATGTGTCC  | Triciccier        | CTGTTTGAC:1'  |
|       | AGTGCATACT   |              | TTCATGGGAG   | Transcher             | ACAM TACAT                 | TANTGAMATA   | Trick CACAT               | CCICTTACAC  | TTTTTCATAC        | ATTRICCCAN:   |
|       | TCACGTATGA   | GATACAGTAN   | ANGTACCUTE   | ALCANACIAS.           | 14:11:4:11:17              | -            |                           |             |                   | A             |
|       | AATAAAGAAT   | GCAAACACAA   | ATCHTTCAAC   | CACAMATAN             | TTCANTITT'A<br>AAGTIVAACCT | CTTITAAAGT   | ACHICATIFITT<br>TCAGTANAA | CAITICAGTAG | TATAGCCCCA        | CCACCACATA    |
|       | GCTTATACAG   |              | CTTAATCAA    | CTCACAGAAC            | CCTACTATTC                 | AACCTACCAC   | CHCCCTCCCA                | ACACACAGAG  | TACACAGTCC        | TTTC:ICCC: YE |
|       | CONATATOTC   | -            | GAATTACTIT   | GAGTOTOTTO            | GGATCATAAG                 | THANACACTO   | GACCCAGGGT                | rendictor   | ATCTCTCAGG        | ANAGAGGG      |
|       | GCTGGCCTTA   | ANANGCATCA   | TATCATGGGT   | AACAGACATA            | THETTAGGTG                 | TTATATTCCA   | CACCETTACE                | TOTOGRAGOCA | AACGCTCATC        | AGTONTATT     |
|       | CGACCGGAAT   | TITTCGTAGT   | ATAGTACCCA   | TTGTCTGTAT            | NA:NATCCAC                 | AATATAAGGT   | <b>OTFICE ANALIG</b>      | ACAGCTCGGT  | TTCCCAGTAG        | TCACTATA      |
|       | ATAMCTECE    |              | ACTTANGTTC   | ATGTCCCTGT            | ניבאניביונים               |              | TCCTCTCCAA                | CANCICATING | CTTAACGGGC        | ממבנואימנואי: |
|       | TATITIOAGGG  | GCCCGTCGAG   | TOANTTCAAG   | TACAGCGACA            | CASTACGACGAC               | 100000       | M. CM. ANGOLIS            | owncocome   |                   |               |
|       |              |              |              |                       |                            |              |                           |             | !                 | !             |
|       | ANOTECACOC   | : CTACATOGGG | GTAGAGTCAT   | AATCGTGCAT            | CAGGATAGGG                 |              | CCAGCAGCGC                | GCGANTAAAC  | TOCHOCOCC         | OCCUCATOR     |
|       | TTCAGGTOCO   | 3 GATGTACCCC | CATCTCAGTA   | Tracacca              | GTCCTATCCC                 | GCCACCACGA   | COTCOTCOCG                | COCTTATTTO  | ACCACGCCOG        | COCCONOCCA    |
|       | Pari         |              |              |                       |                            |              |                           |             |                   | -             |
|       | CUTTALACTORA | TACAACATOO   | CAGTGGTCTC   | CTCAGCGATG            | ATTEGEACEG                 | CCCGCAGCAT   | AAGGCCCCTT                | GTCCTCCGGG  | CACAGCAGCG        | CACCCTOA!     |
|       | GOACOTCCTT   |              | GTCACCAGAG   | GAGTCGCTAC            | TAAGCGTCAC                 | GGGCGTCGTA   | TTCCGCCCAA                | CAGGAGGCCC  | Grorestese        | GTCCCACT***   |
|       |              |              | Patt         |                       |                            |              |                           |             |                   |               |
|       | TCACTTAAAT   | r CAGCACAGTA | ACTOCAGOAC   | ACCACCACAA            | TATTEMENA                  | AATCCCACAG   | TGCAAGGCGC                | TGTATCCAAA  | CCTCATGGCG        | OCCACACA      |
|       | AGTGAATTTA   |              | TGACGTCGTG   | TCCTGGTGTT            | ATAACAAGTT                 | TTAGGGTGTC   | ACGITTCCGCG               | ACATAGGITT  | CONGINCOCC        | CCCTGGTGTC    |
|       | AACCCACGTG   | 3 GCCATCATAC | CACAAGCGCA   | CCTAGATTAA            | GTGGCGACCC                 | CTCATABACA   | CRCTGGACAT                | AAACATTACC  | TCTTTTGGCA        | TOPTGTAATT    |
|       | THOOOTIGEAC  | _            | GIGTICGCGT   |                       | CACCGCTGGG                 | GAGTATTIOT   | GCGACCTGTA                | TITICTAATGG | AGNAAACCOT        | ACAACATTAA    |
|       |              | Kpri         |              |                       |                            |              |                           |             |                   | . Anna        |
|       |              | CCCTACCATA   | TAAACCTCTG   | TARACCICIS ATTARACATE | GCGCCATCCA                 | CCACCATCCT   | MACCAGCTG                 | GCCAAAACCT  | 2002202220        | TATACACTXX.   |
|       | GTGGTGGAGG   | _            | ATTTEGRENCIA | TAATTIGTAC            | CGCGGTAGGT                 | GGTGGTAGGA   | TITICKTICGAC              | COCTITINGGA | COCOCCCCC         | ATATGTGAC:    |
|       | Pad          |              |              |                       |                            |              |                           | ErofilV     |                   | ,             |
|       | ACCESTANCES  | T CALTROPACA | ATGACAGTOG   | AGAGGCCCAGG           | ACTUTATACC                 | ATCGATCATC   | ATRICTEGICA               | TGATATCAAT  | GTTGGCACAA        | CACAGGCACA    |
|       | Thermal      |              | TACTGTCACC   |                       |                            | TACCTAGTAG   | TACGAGCAGT                | ACTATAGETEA | CAACCGTGTT        | GTGTCCGTGT    |
|       |              |              |              |                       |                            | •            |                           |             |                   | Pell          |
| 33301 | COTOCATACA   | A CTTCCTCAGG | ATTACAAGCT   | CCTCCCGCGT            | TAGAACCATA                 | TCCCAGGGAA   | CAACCCATTC                | CTGAATCAGC  | <b>OTAMATCCCA</b> | CACTRICAGG    |
|       | GCACGTATOT   | T DANGGACTCC | TAATGTTCCA   | GGARGGCGCA            | ATCTINGTAT                 | AGGCTCCCTT   | CTTCCCTAAG                | GACTTAGICO  | CATTIAGGGT        | GTCACCTCC     |
| 33401 | AAGACCTCGC   | C ACGTAACTCA | COTTGTGCAT   | TGTCAAGTG             | TTACATTCCG                 | GCAGCAGCGG   | ATGATCCTCC                | AGTATOGTAG  | CCCCCCTTTC        | TOTOTOAN      |
|       | TTCTOGAGCO   | 0 TOCATTOAGT | GCAACACGTA   | ACAGITITCAC           | ANTOTAAGCC                 | نىدىدىدى     | TACTAGGAGG                | TCATACCATC  | OCOCCCANAG        | ACAGAGETET F  |
| 33501 | CICAGGTACAC  | C GATCCCTACT | GTACCAAGTG   | CGCCGAGACA            | ACCCAGATEG                 | ACHICACHUCAL | AGTGTCATG                 | CANATEGAAC  | CCCCCACCTA        | GICATAFIF.    |
|       | CCTCCATCTO   |              |              |                       | TEGETETAGE                 | ACAACCAGCA   | TCACAGTACG                | GITTACCTIO  | CGGCCTGCAT        | CAGTATAAV.    |

|          |       |  |                          |   | - Secondon   |  |                   |                |                          |                          |                |
|----------|-------|--|--------------------------|---|--|--|-------------------|----------------|--------------------------|--------------------------|----------------|
| 33       | 33601 | CTCAACCAAA   |                          | GOCGTGACAA                              | ACAGATETEC   | OTCINCAGE  | hyderecentry (    | GATCGCTCTG     | TCTAGTAGTT<br>ACATCATCAA | GTARTATATC<br>CATCATATAG | CACTCTCTTA     |
|          |       | GACIFICOTITI   | TAGICCACIAC              | ברניכאי חידוד                           | וכון יאייינים  |  |                   |                |                          |                          |                |
| 33       | 33701 | AAGCATCCAG   | <b>OCCCCCCTTG</b>        | CCTTCGGGTT                              | CTATGTAAAC   | TCCTTCATCC   | . טיינינינינעניני | TICATANCATIC : | CACCACCGCA               | GANTAAGCCA               | CACCCAGCC      |
|          |       | TTCOTAGGTC   | COCCOCCCAC               | CGAAGCCCAA                              | GATACATTE  | AGGAAGTACG   | COUCCACGGG        | ACTATTOTAG     | GTCGTGGCGT               | CITATICGGE               | GTGGGTCGGT     |
| _        | 1001  | Perinana and and and and and and and and and   | _                        | ACTUACACAC                              | CERTAGORICO  | GGAAGAGGTG   | CAACAACCAT        | CHARLE PRINTER | TTATTCCAAA               | AGATTATCCA               | ANACCTI YAA    |
| י        |       | Trachterera  |                          | TCAGTGTGTG                              | CCCTCCTCCC   | CUTTCTURAC   | CTTCTTGGTA        | CHANAMANA      | AATAAGGTTT               | TCTAATAGGT               | TITICGAGITT    |
|          |       | Bgfff  |                          |   |  |  |                   |                |                          |                          |                |
| -        |       |  | A COLUMN DE COLUMN DE DE |   | ماديدون فالمتادون  | Transfer BAACT   | CTACACCCAA        | AGNACAGATA     | ATCOCATITO               | TAAGATGTTO               | CACANTGGCT     |
| 7        | 70666 | AIGAGOTCI  |                          | במכנוניונים                             | NCCC NCCC  | ACCACTTEGA   |                   | TESTETAT       | TACCGTAMAC               | ATTCTACAAC               | GTOTTIACCGF    |
| 7        |       | IACIICIAGA   |                          | Canada and Canada                       | TOO BOOM BE  | CCTTBBBCCC   | TTCACCATTCA       | ATEMPORAL      | TAVACAPTICE              | ACCACCTTCA               | ACCATOCCCA     |
| Ť        | 34001 | TCCAAAAGGC   |                          | CACCACCAGG                              | ACCIPICANTER   | CCCATTENDO   | AACTUCUT          | TAGAGGAGAT     | ATTTOTANGO               | TCGTGGAAGT               | TCGTACCCCT     |
|          |       | AGGITTICCO   |                          | CICCACCITC                              | אנכוייניי  | 200111000  | Maicence          |                |                          |                          |                |
| 34       | 34101 | ANTANTICIC   |                          | CTTCTCAATA                              | TATCTCTAAG   | CAMATCCCGA   | ATATTAAGTC        | CGGCCATTGT     | AAAAANCIGC               | אככשבשנוסנים             | CCICCACCII     |
|          |       | TTATTAAGAG   | TAGAGCGGTG               | GAAGAGTTAT                              | ATAGAGATTC   | GTTTACCCCT   | TATAATTCAG        | GCCGGTAACA     | TTTTTAGACG               | ASSICIONE                | Newson Control |
| 34       | 34201 | CAGCCTCAAO   | CAGCGAATCA               | TGATTGCAAA                              | AATTCAGGTT   | CCTCACAGAC   | CTCTATAAGA        | TTCAAAAGCO     | GAACATTAAC               | AAAAATACCO               | CGATCCCGTA     |
|          |       | GTCOGAGTTC   | OTCOCTFAGT               | ACTAACCITY                              | TTAAGTCCAA   | GGAGTOTCTG   | GACATATTCT        | MOTITICOC      | CITICITANITIC            | THITTATOGC               | GCTAGGGCA1     |
| 72       | 10191 |  | CAGGGGCCAGC              | TGAACATAAT                              | CGTGCAGGTC   | TGCACGGACC   | AGCCCCCCA         | כיווככככככ     | ACCANCCATG               | ACANAGGAC                | CCACACTOAT     |
| •        | 1     | CCAGGGAAACC  | _                        | ACTITICITATITA                          | GCACGTCCAG   | ACCITGCCTGG  | TCGCGCCGGT        | GAAGGGGCGG     | TCCTTGGTAC               | TOTALING                 | GGTGTGACTA     |
|          |       |  |                          |   |  | Ŧ  |                   |                |                          |                          |                |
| 2        | 14401 | The Strange of the St | ATACTORGAG               | CENTRICE AND                            | CACCCTACCC   | CCGATGTANG   | CTTGTTGCAT        | GGGCGCGAT      | ATAAAATGCA               | Aggreeneer               | CAAAAAATCI     |
| <b>`</b> | T 0 P | ATACHTERIOR  | TATGAGCTC                | GATACGATTG                              | GTCGCATCGG   | COCTACATTC   | GAACAACGTA        | CCCGCCGCTA     | TATTITIACGE              | TCCACGACGA               | GTTTTTAG       |
| i        | ,     | 1  |                          | 40400404                                |  | LINUTE WARE  | ATABACCIAG        | GTAAGCTCCG     | GAACCACCAC               | AGAAAAAGAC               | ACCATHITIC     |
| ň        | 345UI | GGCAAAGCCT   |                          | MINISTER COLUMN                         | ATOMOTION OF THE PROPERTY OF T | OCTACE PORTER  | PATTHICKER        | CATTCGAGGC     | CITIGOTOOTO              | TCPTTTTCTG               | TOOTANAAAG     |
| . `      |       | ברפון ונפט   |                          | 101111111111111111111111111111111111111 | TION THE TOTAL   | Description of the same of the | A A A C A TOTAL   | AACATTAGAA     | GIT WENT TIER            | CAACAGGAAA               | AACAACCC711    |
| ň        | 34601 | TCTCAMCAT  | _                        | TICICCATAA                              | ACACAMATA  | AVATARCANI   | TTTTCTABAT        | THETABATETA    | CGCACAGAAT               | GETOTCCTET               | THEFTOCOAN     |
|          |       | AGACITITICITA  | CAGACOCCCA               | AAGACGTATT                              | TGTGTGTTA  |  | 100000            |                |                          |                          |                |
| ň        | 34701 | ATAMOCATAA   | GACCOCACTAC              | OCCCATGCCG                              | GCGTGACCGT   | NAVABACTG  | GTCACCOTGA        | TTANAMGCA      | CCACCCACAG               | Chechesis                | ATC TCCCTC     |
|          |       | TATTCGTATT   | · CTGCCTGATO             | CCCCTACGC                               | CCCACTCCCA   | TITITITIENC  | CAGTGGCACT        | AATTTTCGT      | GGTGGCTGTC               | GALEGALACEAG             | ויי.אפיירור    |
| ř        | 34801 | TCATAATOTA   | AGACTCOGTA               | AACACATCAG                              | GTTGATTICAC  | ATCGGTCAGE   | GCTAANNGC         | GACCGANATA     | GCCCCCCCCCCA             | ATACATACCC               | CHENCHARICETAG |
| •        |       | AGTATTACAT   | -                        | TTGTCTAGTC                              | CAACTAAGTG   | TAGCCAGTCA   | CGATTTTCG         | CTGGCTTTAT     | COOCCCCCT                | TATOTATOCG               | CONCOCATO      |
| Ē        | 14011 | SCACARCATE   | ACACCCCC                 | TACCACCTAT                              | AACAAAATTA   | ATAGGAGAGA   | ANANCACATA        | AACACCTGAA     | ANACCCTCCT               | GCCTAGGCAA               | AATAGCACCC     |
|          | 100   | Tr-representative  | · .                      | ATCCTCCATA                              | THGTTTTANT   | TATCCTCTCT   | TTTTICHCTAT       | TTGTGGACTT     | TTTCCGAGGA               | COGATCCCTT               | TTATCGTGGG     |
| ř        | 10035 |  |                          | rancemen                                | ACAGCGGCAG   | CCATAACAGT   | CAGCCTTACC        | ACTANAMAG      | AAAACCTATT               | MANANACAC                | CACTCRACA:     |
| •        | 1000  | Actorican  | _                        | GTCGCGAAGG                              | TGTCGCCGTC   | GGTATTGTCA   | CTCCCAATCG        | TCATTTTTC      | TTTTGGATAA               | THEFTIGIE                | GTOAGCTGTN3    |
| ř        | 10125 |  |                          | CAGTGTAAAA                              | AAGGGCCAAG   | TGCAGAGCGA   | CTATATATAG        | GACTAAAAA      | TGACGTAACG               | <b>GTTAAAGTCC</b>        | ACAAAAAAAA     |
| 1        | 1010  | CCOTOOTCGA   |                          | GTCACATTTT                              | TYCCCGGTTC   | ACGICICACCT  | CATATATATC        | CTCATTITIT     | ACTICCATTICC             | CAATITICAGG              | 1CTTTTTGI      |
| -        | 15201 | CCCAGAAAC  | COCACOCOAA               | CCTACGCCCA                              | GANACGANAG   | CCANNANCC  | CACAACTTCC        | TCAAATCGTC     | ACTROCOPTE               | TCCCACGTTA               | CGITCACTITICC  |
| )        |       | GGCTCTTTTG   | -                        | CGATGCCCGT                              | CTTRCCTTC  | GCTTTTTGG  | GTCTTGAAGG        | AGTITAGGAG     | TOMOGENM                 | AGGGTGCAAT               | GCAGTGAAGG     |

Figure 15V

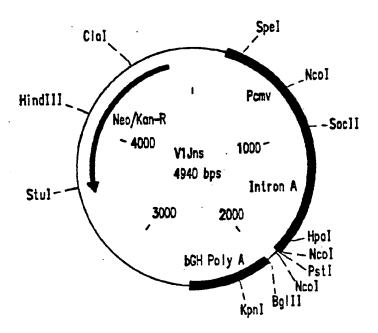
| 16101 | A Shipping B.O.B.        | -441441444               | Truck Branch             | CATTER   | W.L.C.C.C.                | AACCTACCTIC ACCCGCCCG     |                           | Trechoge                 | CCGCCACG                 | TCACAAACTC                 |
|-------|--------------------------|--------------------------|--------------------------|--|---------------------------|---------------------------|---------------------------|--------------------------|--------------------------|----------------------------|
| 1000  | GTANANTICT               | TITIGATETTA              |                          |  | MCCCCCCCATT               | TYCGATCCAG TYGGCCGGCC     |                           |                          | GCCCCGOTGC               | ACTOTITIONS                |
| •     |                          |                          |                          |  |                           | Pac                       | el<br>Ecnfil              | _                        |                          |                            |
| 35401 | CACCCCCTCA               | TTATCATATT               | GGCTTCAATC               | CAAAATAAGG                                     | TATATTATTG                | TACTACANT                 | TTAAGAATTC                | GGATCTGCGA               | CCCCAGCTO                | GATEGECTT:                 |
| 35501 | CCCATTATOA               | THETTERE                 | 1100000000               |  |                           |                           | TCCAGGCAGG                |                          | CCATCAGGGA               | CACCTICAAG                 |
| 1     | GCOTAATACT               | AAGAAGAGCG               | AAGGCCGCCG               |  | GGCGCAACGE                |                           | AGGICCOICL                |                          | A PARAMETER I            | CITAC ACTACL               |
| 35601 | CCCACCAAAA               | GGCCAGGAAC               | CCTTAAAAAGG              | CCCCOTTCCT                                     | CCCCANANG                 | CATAGGCTCC                | GCCCCCCTGA                | CONGCATCAC               | TTTTAGCTO                | CCAOTTCAGT                 |
| 35701 | DAGOTOCCOA               |                          | GACTATAAAG               | ATACCAGGCG                                     | TTTCCCCCTC                | GAAGCTCCCT                | CONGCOCICT                | CCTUTTCCGA               | CCCTGCCGCT               | TACCCCATAC                 |
| 35801 | CTOTOCOCCT               |                          | GGGAAGCGTG               | -  | ATAGCTCACG                |                           | CTCANTTCGG                |                          | Trochecano               | CTCCCTCT                   |
| 35901 | TREACTORACE              | CCCCGTTCAG               | CCCTTCCCAC               | GCCCCTTATC                                     | CGGFAACTAT                | COTCITICAGE               | CCAACCCAGT                |                          | TTATCGCCAC               | TOOCAGCAG                  |
|       | ACCITICA                 | _                        |                          | CGCGGMATAG                                     | GCCAFFGATA                |                           | GGTTGGGCCA                | reterocro                | AATAGCCCTO               | ACCGTCGTC                  |
| 36001 | CACTOGTAAC               | AGGATTAGCA               | CTCGCTCCAT               | TOTAGGCGGT<br>ACATCCGCCA                       | GCTACACAGT                | TCTTCAAGTG<br>AGAACTTCAC  | GTGGCCTAAC                | TACCOCTACA               | CTAGANGGAC<br>GATCTTCCTG | AGTATTTGGF<br>TCATAAACCA   |
| 36101 | ATCTOCGCTC<br>TAGACGCGAG |                          |                          | CCTTTTTCTC                                     | TTCSTAGGTC                | THEATTCEGGE               | AVACAAACCA                | CCCCTGGTAG               | CCCACCAAAA               | TTTGTTTGC.<br>AAACAAACGT   |
| 36201 | AGENGEAGAT               | -                        |                          |  | TCCTFTGATC<br>AGGAAACTAG  | TITICINGG                 | GGTCTGACGC                | TCAGTGGAAC<br>AGTCACCTTG | GAMACTCAC                | GTTANGGGA"<br>CAATTCCCTA   |
| 36301 | AAACCAGTAC               |                          |                          |  | CTTTTAAATC                | AATCTAAAGT                | ATATATGAGT<br>TATATACTCA  | ANACTTEGOTC              | TOACAGTTAC<br>ACTOTCAATO | CNATCCTFIN                 |
| 36401 | TCAGTGAGGC               |                          | OCGANCTOTC<br>CGCTAGACAG |  | ATCCATAGIT                | GCCTOACTCC                | CCGTCGTGTA                | GATAACTACG               | ATACGGGAGG<br>TATGCCCTCC | GCTTACCATY:<br>CGAATCGTAG  |
| 36501 | TOCCCCAOT                |                          | TACCCCGAGA               | CCCACOCTCA                                     | CCGGCTCCAG                | ATTTATCAGE<br>TAAATAGTCG  | AATAAACCAG<br>TTATTTGGTC  | CCACCCCAA                | GGGCCGAGCG               | CACIANGTOOT                |
| 36601 | CCTOCAACTT               | FATCCOCCTC<br>ATACATARAG | CATCCAGTCT               | TAATTAATTGIT                                   | CCCCCCTTCG                | TAGAGTAAGT                | AGTTCGCCAG<br>TCAAGCGGTC  | TTAATAGTTT               | GCGCAACGTT               | GANCGGTAAC                 |
| 36701 | CTACAGGCAT               |                          |                          |  | TTCATTCAGC                | TCCGGTTCCC<br>AGGCCAAGGG  | AACGATCAAG                | GCGAGTTACA               | TGATCCCCCA<br>ACTAGGGGGT | TOTTOTICAN<br>ACAACACOTT   |
|       |                          |                          | Pytt                     | 11<br>******                                   |                           |                           |                           |                          |                          | !                          |
| 36801 | MAAGCGGTT                | AGCTCCTTCO               |                          | GTCCTCCOAT CGTTGTCAGA<br>CAGGAGGCTA GCAACAGTCT | ACTAAGTTOG<br>TCATTCAACC  | CCCCAGTGTT                | ATCACTCATG<br>TAGTIGAGTAC | CANTACCOTC               | CACTGCATAA               | ANGAGAATGA                 |
| 36901 | GTCATGCCAT               | CCGTAAGATO               | CITITICITATA GAAAAGACAC  | ACTOGTGAGT<br>TGACCACTCA                       | ACTY MACCAA<br>TGAGTTGGTT | CACTATACTOA<br>CACTAAGACT | GAATAGTGTA                | TOTOCCCACC<br>ACCCCCCTGG | GAGTTTGCTCT              | Tracce (Age of Acceptance) |

figure 15W

### DMRKAdSgag MERGB

| CNICACCODA FANTACCECO CCACATAGEA GAACTITAAA AGRECICATE AFTGAAAAA GATCHIGGO GGGAAAACTE TEAAGGAETE TACRECISTET G<br>GTIGTGCCCT ATPATGGCG GGTOTATCE CPTCAAATTI TCACAGTAG TAACCTITTG CAAGAAGCC GOCTITIGA AGTCCTAGA AFTGCGAAAA GAAAACAA AAACCAAAA CCACAGGATAG GAAAAACAG AAACCAAAAA CTCAAGGTAA GAAAATTAGTC TACAGGAGA GAAAAACAA AAACCAAAAAAAAA CTCAAAAATTAGCC TACAGTAGAA AAACCAACT GAAAATTAGAAAAAAAAAA | OCCOCADADA AGGGATTAG GGCGACACGG AAATGTTGAA TACTCATAGT CTTCCTTTTT CAATATTAT GAAGGATTA TCAGGGTTAT TCAGGGTTAT TCAGGGTTAT TCAGGGTTAT TCAGGGTTAT TCAGGGTTAT TCAGGGTTAT TCAGGGTTAT TCAGGGTTAT TCAGGGTTACAT ACAGGGTTACAT AGGGGTTATC CCCGAAAAGT GCCACGGAA CTTAGGAAA CTTAGAAAA ATAAAAAAAA ATATAGAAAA ATATAGAAAA AGGGGTTATC CTCAAAAAAAAAA | EGOTH  BATTALL  BATTALL  BATTALL  CATORCATTA ACCTATABADA ATARGEGINT CACGARGEGE THTCGTCTTC ANGANTIEGA TECGANITICT THAT (SEQ ID NO: 27)  GIACTOTABAT INGATAITIT TATECECATA GIGCTCCHOG AAAGARAAG TICTTAMACT AGGCTTAMGA ATTA (SEQ ID NO: 28) |
|---|---|--|
| GTTCTTCGGG GCC CAACAAGCCC CCC CACCAAGCGTT TCT GTCGTAGGGTA AGA   | CANTATTATT GAN<br>GTTATAATAA CTT<br>CCCGAAAAGT GCC<br>GGGCTTTTCA CGC  | Feetti  TUCGAATTCT TAA AGGCTTAAGA ATT  |
| TANCCTITIGO CTITITACTITI  | CANCITITE<br>GANGIAMAA<br>CIKACATITIC<br>GCGIGTAMAG   | Bamfill<br>MAGANTTEGA TO   |
| AGTESCECATE TEACGACTAG TETTEACTAGE AGAACTECTA   | TACTCATACT ATTACTTATES AGGRETTTCC   | TITEGICITE   |
| GAACTTTAAA<br>CTTCAAATTT<br>ACTCAACTGA<br>TGGGTTGACT  | AAATGTTGAA<br>TTTACAACTT<br>ATAAACAAAT  | CACGAGGCCC   |
| CCACATAGGA<br>GGTGTATGGT<br>CCACTCGTGC<br>GGTGAGGAGG  | GGCGACACGG<br>CCCCTCTTCCC<br>ATTTAGAAAA<br>TAAATCTTTTT  | ATACGCGTAT<br>TATCCGCATA   |
| TAATACCGCG<br>ATTATGGCGC<br>TCGATGTAAC<br>AGCTACATTG  | AGGGNATAAG<br>TCCCTTATTC<br>ATTTGAATGT<br>TAAACTTACA  | ACCTATAAAA<br>TGGATATTTT   |
| CAACACGGA<br>GTTGTGCCCT<br>GAGATCCAGT<br>CTCTAGGTCA   | GCGCGAAAA<br>CGGCGTTTTT<br>GCGGATACAT<br>CGCCTATGTA   | <b>CATGACATTA</b><br>GTACTGTAAT  |
| 37001   | 37201   | 37401  |

Figure 15X



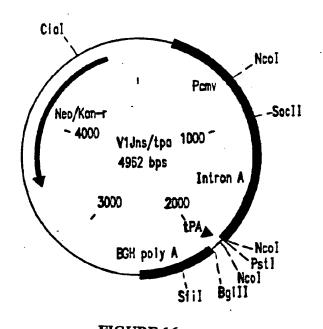


FIGURE 16

GCAGTGGCCCCTGACTGAGGAGAAGATCAAGGCCCTGCTGGAAATCTGCACTGAGATGGAGAAGGAGGGCAAAATCTCCA sGinTrpProLeuThrGluCluLyslleLysAloLeuVolGluIleCysThrGluMeLGluLysGluGlyLysIleSerL 30 40 50

AGATTGGCCCCGAGAACCCCTACAACACCCCTGTGTTTGCCATCAAGAAGAAGAAGGACTCCACCAAGTGGAGGAAGCTGGTG
ysileGiyProGluAsnProTyrAsnThrProVolPheAlolleLysLysLysAspSerThrLysTrpArgLysLeuVol
60 70

GACTTCAGGGAGCTGAACAAGAGGACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCCACCCCGCTGGCCTGAAGAA AspPheArgGluLeuAsnLysArgThrGlnAspPheTrpGluVolGlnLeuGlyIleProHisProAloGlyLeuLysLy 80 90 100

GAAGAAGTCTGTGACTGTGCTGGCTGTGCGGGATGCCTACTTCTCTGTGCCCCTGGATGAGGACTTCAGGAAGTACACTG slyslysSerVoiThrVoiLeu<u>Alo</u>VoiGiyAspAloTyrPheSerVoiProLeuAspGluAspPheArgLysTyrThrA 110 120 130

CCTTCACCATCCCCTCCATCAACAATGAGACCCCTGGCATCAGGTACCAGTACAATGTGCTGCCCCAGGGCTGGAAGGGC
IoPheTnrlleProSerileAsnAsnGluThrProGly!leArgTyrGinTyrAsnVoiLeuProGinGlyTrpLysGly
140 150

TCCCCTGCCATCTTCCAGTCCTCCATGACCAAGATCCTGGAGCCCTTCAGGAAGCAGAACCCTGACATTGTGATCTACCA SerProAioliePheGinSerSerMetThrLyslieLeuGiuProPheArgLysGinAsnProAsplieVoilleTyrGI 160 170 180

GTACATGCCTGCCCTGTATGTGGGCTCTGACCTGGAGATTGGGCAGCACGACCAAGATTGAGGAGCTGAGGCAGCACC
nTyrMeialoaloleuTyrVoiGiySerAspLeuGiulleGiyGinHisArgThrLysIleGiuGiuLeuArgGinHisL
190 200 210

TGCTGAGGTGGGGCCTGACCACCCCTGACAAGAAGCACCAGAAGGAGCCCCCCTTCCTGTGGATGGGCTATGAGCTGCAC euleuArgTrpGlyLeuThrThrProAspLysLysHisGlnLysGluProProPheLeuTrpMetGlyTyrGluLeuHis 220 230

CCCGACAGTGGACTGTGCACCCATTGTGCTGCCTGAGAAGGACTCCTGGACTGTGAATGACATCCAGAAGCTGGTGGG
ProAspLysTrpThrVoIGInProIIeVoILeuProGluLysAspSerTrpThrVoIAsnAspIIeGInLysLeuVoIGI
240 250 260

CAAGCTGAACTGCGCCTCCCAAATCTACCCTGGCATCAAGGTGAGGCAGCTGTGCAAGCTGCTGAGGGCACCAAGGCCCC
yLysLeuAsnTrpAioSerGinlieTyrProGiylieLysVoiArgGinLeuCysLysLeuLeuArgGiyThrLysAioL
270 280 290

### FIGURE 17A

TGACTGAGGTGATCCCCCTGACTGAGGAGGCTGAGCTGGAGCTGGAGCAGACAGGGAGATCCTGAAGGAGCCTGTGCAT EUThrGluVollleProLeuThrGluGluAloGluLeuGluLeuAloGluAsnArgGluIleLeuLysGluProVolHis 300 310

GGGGTGTACTATGACCCCTCCAAGGACCTGATTGCTGAGATCCAGAAGCAGGGCCAGGGCCAGTGGACCTACCAAATCTA GlyVolTyrTyrAspProSerLysAspLeulleAloGiulleGinLysGinGlyGinGlyGinTrpThrTyrGinlleTy 320 330 340

CCAGGAGCCCTTCAAGAACCTGAAGACTGGCAAGTATGCCAGGATGAGGGGGGCCCCACACCAATGATGTGAAGCAGCTGA rGInGIuProPheLysAsnLeuLysThrGiyLysTyrAloArgMetArgGlyAloHisThrAsnAspVolLysGInLeuT 350 360 370

CTGAGGCTGTGCAGAAGATCACCACTGAGTCCATTGTGATCTGGGGCAAGACCCCCAAGTTCAAGCTGCCCATCCAGAAG hrGluAloVoiGinLyslieThrThrGluSerlieVoilieTrpGlyLysThrProLysPheLysLeuProlleGinLys 380 390

GGTGAAGCTGTGGTACCAGCTGGAGAAGGAGCCCATTGTGGGGGCTGAGACCTTCTATGTGGCTGGGGCTGCCAACAGGG uVolLysleuTrpTyrG1nLeuG1uLysG1uProl1eVo1G1yA1oG1uThrPheTyrVo1A1oG1yA1oAsnArgG 430 440 450

AAGACTGCCCTCCAGGCCATCTACCTGGCCCTCCAGGACTCTGGCCTGGAGGTGAACATTGTGACTGCCTCCCAGTATGC
LysThrAloLeuGinAlolleTyrLeuAloLeuGinAspSerGiyLeuGiuVolAsnIieVolThrAloSerGinTyrAi
480
490
500

CCTGGGCATCATCCAGGCCCAGCCTGATCAGTCTGAGCTCTGAGCTGGTGAACCAGATCATTGAGCAGCTGATCAAGAAGG ©LeuGiylielieGinAloGinProAspGinSerGiuSerGiuLeuVolAsnGinIlelieGiuGinLeuIleLysLysG 510 520 530

AGAAGGTGTACCTGGCCTGGCCTGCCCACAAGGCCATTGGGGGCAATGAGCAGGTGGACAAGCTGGTGTCTGCTGGC
IULysVolTyrLeuAloTrpVolProAloHisLysGlylleGlyGlyAsnGluGlnVolAspLysLeuVolSerAloGly
550

ATCAGGAAGGTGCTGTTCCTGGATGGCATTGACAAGGCCCCAGGATGAGCATGAGAAGTACCACTCCAACTGGAGGGCCTAT

1 leAr gl ysVoil euPheleuAspGiylleAspLysAioGinAspGluHisGluLysTyrHisSerAsnTrpArgAioMe
560 570 580

### FIGURE 17B

CGCCTCTGACTTCAACCTGCCCCCTGTGGTGCCTAAGGAGATTGTGGCCTCCTGTGACAAGTGCCAGCTGAAGGGGGAGG tAloSerAspPheAsnLeuProProVolVolAloLysGlulleVolAloSerCysAspLysCysGlnLeuLysGlyGluA 590 600 610

GCTGTGCATGTGGCCTCCGGCTACATTGAGGCTGAGCTGATCCCTGCTGAGACAGGCCAGGAGACTGCCTACTTCCTGCT AlovolHisVolAloSerGlyTyrIleGluAloGluVollleProAloGluThrGlyGlnGluThrAloTyrPheLeuLe 640 650 660

GAAGCTGGCTGGCAGGTGGCCTGTGAAGACCATCCACACTGCCAATGGCTCCAACTTCACTGGGGCCACAGTGAGGGCTG uLysLeuAloGlyArgTrpProVolLysThrlleHisThrAloAsnGlySerAsnPheThrGlyAloThrVolArgAloA 670 680 690

CCTGCTGGTGGCCTGGCATCAAGCAGGAGTTTGGCATCCCCTACAACCCCCAGCGGGTGGTGGCCTCCATGAAC IoCysTrpTrpAloGlylleLysGlnGluPheGlylleProTyrAsnProGlnSerGlnGlyVolVolAloSerMelAsn 700 710

AAGGAGCTGAAGAAGATCATTGGGCAGGTGAGGGACCAGGCTGAGCACCTGAAGACAGCTGTGCAGATGGCTGTGTCAT LysGluLeuLysLyslielleGlyGlnVolArgAspGlnAloGluHisLeuLysThrAloVolGlnMetAloVolPheli 720 730 740

CCACAACTICAAGAGGAAGGGGGCATCGGGGGCTACTCCGCTGGGGAGAGGATTGTGGACATCATTGCCACAGACATCC eHisAsnPheLysArglysGlyGlylleGlyGlyTyrSerAloGlyGluArglleVolAsplleIleAloThrAsplleG 750 760 770

AGACCAAGGAGCTCCAGAAGCAGATCACCAAGATCCAGAACTTCAGGGTGTACTACAGGGACTCCAGGAACCCCCTGTGG
InThrLysGluLeuGInLysGInileThrLyslieGInAsnPheArgVolTyrTyrArgAspSerArgAsnProLeuTrp
780 790

AAGGCCCCTGCCAAGCTGCTGTGGAAGGGGGAGGGGGTGTGGTGATCCAGGACAACTCTGACATCAAGCTGGTGCCCAG LysGtyProAtalysLeuleuTrpLysGtyGtuGtyAtoVatteGtnAspAsnSerAspIteLysVatVatProAr 800 810 820

AAACCCCCCCCACATC; (SEQ ID NO: 3)

Xx Bg/11 (SEQ ID NO: 4)

FIGURE 17C

GATCACCATGGATGAAGAGAGGCTCTGCTGTGTGCTGCTGTGTGCAGCAGTCTTGCTTTGCC
MetAspAfdMetLysArgGjyLeuCysCysVolLeuLeuLeuCysGjyAloVoiPheVoiSerP
-25
-25

### FIGURE 18

| WT              | - ATG GGT GGC AAG TGG TCA AAA CGT AGT GTG CCT GGA TGG TCT -42                                |    |
|-----------------|--|----|
| OPT             | - ÁTG GÓC GÓC ÁÁG TGG TĆC ÁÁG AGG TCC GTG ČĆC GGC TGG TĆC<br>M G G K W S K R S V P G W S -14 |    |
| WT              | - ACT GTA AGG GAA AGA ATG AGA CGA GCT GAG CCA GCA GCA GAT -84                                |    |
| OPT             | - ÁCC GTG AĞG ĞÂĞ ÂĞG ÁTĞ ÂĞG AĞG ĞČC ĞÂĞ ČČC GCC GCC GAC<br>T V R E R M R R A E P A A D -28 |    |
| WT <sup>*</sup> | - AGG GTG AGA CGA ACT GAG CCA GCA GCA GTA GGG GTG GGA GCA -126                               | 5  |
| OPT             | - ÁGG GTG ÁGG ÁGC GÁG CÓC GCC GTG GGC GTG GGC GCC<br>R V R R T E P A A V G V G A -42         |    |
| WT              | - GTA TCT CGA GAC CTG GAA AAA CAT GGA GCA ATC ACA AGT AGC -168                               | 3  |
| OPT             | - GTG TCC AGG GÁC CTG GÁG ÁÁG CÁC GGC GCC ÁTC ÁCC TCC TCC<br>V S R D L E K H G A I T S S -56 |    |
| WT              | - AAT ACA GCA GCT ACC AAT GCT GAT TGT GCC TGG CTA GAA GCA -21                                | )  |
| OPT             | - AÁC ÁCC GCC GCC ÁCC ÁÁC GCC GÁC TGC GCC TGG CTG GÁG GCC<br>N T A A T N A D C A W L E A -70 |    |
| WT .            | - CAA GAG GAT GAG GAA GTG GGT TTT CCA GTC AGA CCT CAG GTA -25                                | 2  |
| OPT             | CAG GAG GAC GAG GAG GTG GGC TTC CCC GTG AGG CCC CAG GTG Q E D E E V G F P V R P Q V -84      |    |
| WT              | - CCT TTA AGA CCA ATG ACT TAC AAG GGA GCT GTA GAT CTT AGC -29                                | 4  |
| OPT             | - CCC CTG ÁGG CCC ÁTG ÁCC TÁC ÁÁG GGC GCC GTG GÁC CTG TCC<br>P L R P M T Y K G A V D L S -98 |    |
| WT              | - CAC TIT TTA AAA GAA AAG GGG GGA CTG GAA GGG CTA ATT CAC -33                                | 6  |
| OPT             | - CAC TTC CTG AAG GAG AAG GGC GGC CTG GAG GGC CTG ATC CAC<br>H F L K E K G G L E G L I H -11 | 2  |
| WT              | - TCA CAG AAA AGA CAA GAT ATC CTT GAT CTG TGG GTC TAC CAC -37                                | 8. |
| OPT             | - TCC CAG AAG AGG CAG GAC ATC CTG GAC CTG TGG GTG TAC CAC S Q K R Q D I L D L W V Y H -12    | 6  |
| WT              | - ACA CAA GGC TAC TTC CCT GAT TGG CAG AAC TAC ACA CCA GGG -42                                | 0  |
| OPT             | - ACC CAG GGC TAC TTC CCC GAC TGG CAG AAC TAC ACC CCC GGC T Q G Y F P D W Q N Y T P G -14    |    |

FIGURE 19A

| WT          | - CCA GGA ATC AGA TTT CCA TTG ACC TTT GGA TGG TGC TTC AAG -462                                |   |
|-------------|---|---|
| OPT         | - CÉC GGC ÁTC ÁGG TTC CÉC CTG ÁCC TTC GGC TGG TGC TTC AAG P G I R F P L T F G W C F K -154    |   |
| ₩T          | - CTA GTA CCA GTT GAG CCA GAA AAG GTA GAA GAG GCC AAT GAA -504                                |   |
| OPT         | - CTG GTG CCC GTG GAG CCC GAG AAG GTG GAG GAG GCC AAC GAG<br>L V P V E P E K V E E A N E -168 |   |
| WT          | - GGA GAG AAC AAC TGC TTG TTA CAC CCT ATG AGC CAG CAT GGG -546                                |   |
| OPT         | GGC GAG AAC AAC TGC CTG CTG CAC CCC ATG TCC CAG CAC GGC  G E N N C L L H P M S Q H G -182     |   |
| WT          | - ATA GAG GAC CCG GAG AAG GAA GTG TTA GAG TGG AGG TTT GAC -588                                |   |
| OPT         | - ATC GAG GAC CCC GAG AAG GAG GTG CTG GAG TGG AGG TTC GAC  I E D P E K E V L E W R F D -196   |   |
| WT .        | - AGC AAG CTA GCA TIT CAT CAC GTG GCC CGA GAG CTG CAT CCG -630                                |   |
| OPT         | - TCC AAG CTG GCC TTC CAC CAC GTG GCC AGG GAG CTG CAC CCC<br>S K L A F H H V A R E L H P -210 |   |
| WT          | - GAG TAC TAC AAG GAC TGC TGA (SEQ ID NO:30) -651   |   |
| <b>O</b> PT | - GAG TAC TAC AAG GAC TGC TAA (contained within SEQID NO:9) E Y Y K D C (SEQID NO:10) -216    | , |

FIGURE 19B

VIJns/nef

SrfI BGIII
CAC CCC GAG TAC TAC AAG GAC TGC TAA AGCCCGGGGAGAICIGCTGTGCC77C7AG77GCCAGC (SEQ 1D NO: 38)
H P E Y Y K D C \* (Contained within SEQ 1D NO: 10:

V1Jns/nef(G2A.LLAA)

*Psti* Catrgastettttelgegicacgtecttgaga<u>tet</u>gecace atg gec ggc ang tgg tec ang agg tec gtg eec . M A G K W S K R S V P

Srff BgIII
. . . CAC CCC GAG TAC TAC AAG GAC TGC TAA AGCCCGGGAGAICIGCTGTGCCTTCTAATTGCCAGC (SEQ ID NO: 39)
H P E Y Y K D C \* (contained within SEQ ID NO:14)

/lJns/tpanef & VlJns/tpanef(LLAA)

Psti Catesorutticiecagicaccotatatatictagaticacc atg gat gca atg ang aga ggg ctc tgc tgt gtg M D A M K R G L C C V

CTG CTG CTG TGT GGA GCA GTC TTC GTT TCG CCC AGC GAG AIC TCC AAG AGG TCC GTG CCC ...

SrfI BOIII

CAC CCC GAG TAC TAC AAG GAC TGC TAA AGCCCGGGGAGAICIGCTGTGCCTICTAGTTGCCAGC (SEQ ID NO: 40)

H P E Y Y K D C \* (contained withon SEQ ID NO: 16)

### FIGURE 20

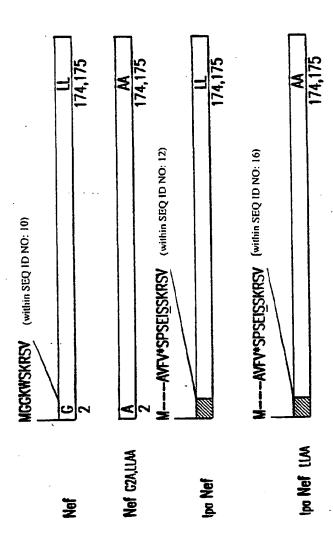


FIGURE 21

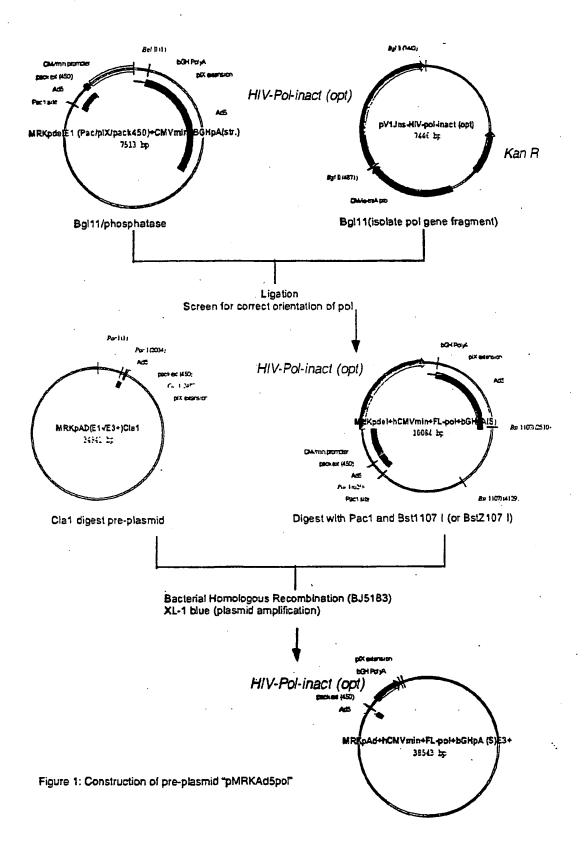


FIGURE 22

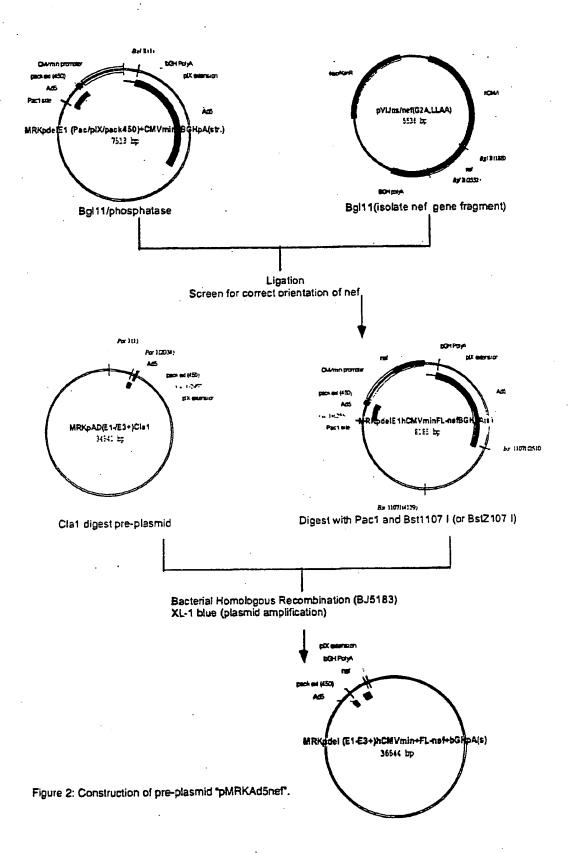


FIGURE 23

Comparison of Clade B vs. Clade C Anti-gag T Cell Responses in Clade B HIV-Infected Subjects

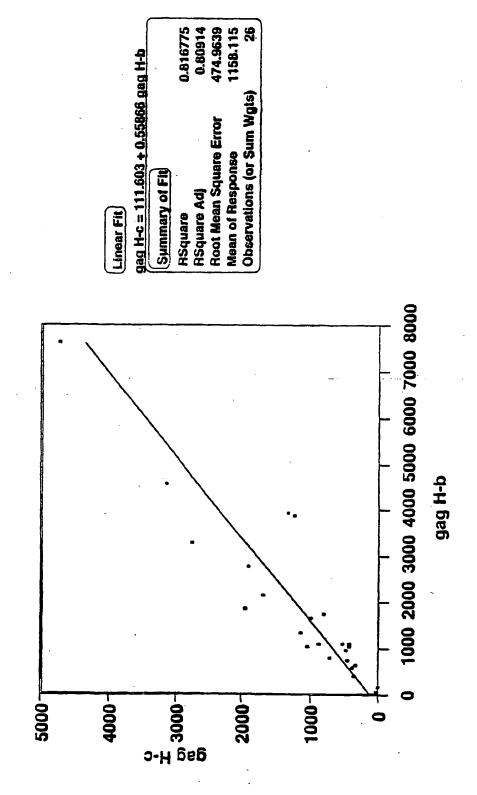
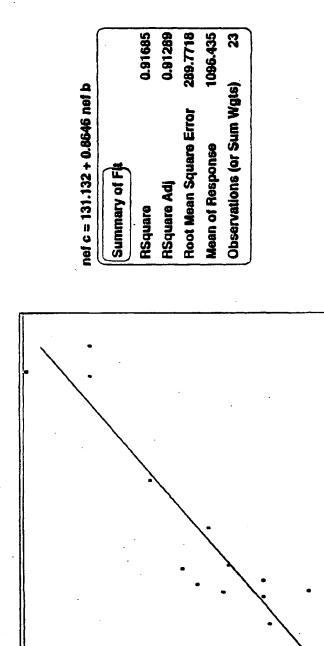


FIGURE 25



nef b

Comparison of Clade B vs. Clade C Anti-nef T Cell

Responses in Clade B HIV-Infected Subjects

3500

3000

2500

net c 2000 1500

1000

-009

### MRKAd5pol MER1062 (MRKAd5 Pre-Adenoviral Vector Containing the IA opt pol Coding Region)

| 1   |   |                        |   | GAAGCCAATA       |                 |
|-----|---|------------------------|---|------------------|-----------------|
|     | GTAGTAGTTA                              | TTATATGGAA             | TAAAACCTAA                              | CTTCGGTTAT       | ACTATTACTC      |
|     |   |                        |   |                  |                 |
| 51  | GGGGTGGAGT                              |                        |   |                  |                 |
|     | CCCCACCTCA                              | AACACTGCAC             | CGCGCCCCCC                              | ACCCTTGCCC       | CGCCCACTGC      |
|     |   |                        |   |                  |                 |
| 101 |   |                        |   | GTGTGGCGGA       |                 |
|     | ATCATCACAC                              | CGCCTTCACA             | CTACAACGTT                              | CACACCGCCT       | TGTGTACATT      |
| 151 | CCCACCCATC                              | <b>TOCOCTA NA NOTA</b> | C \ C \ C \ C \ C \ C \ C \ C \ C \ C \ | GTGTGCGCCG       |                 |
| 131 |   |                        |   | CACACGCGGC       |                 |
|     | COCIOCCIAC                              | ACCUITION              | CIGCHMANC                               | CACACGCGGC       | CACATGTGTC      |
| 201 | GAAGTGACAA                              | TTTTCGCGCG             | GTTTTAGGCG                              | GATGTTGTAG       | ጥል ል ለጥጥጥላጋርጋር  |
|     |   |                        |   | CTACAACATC       |                 |
|     |   |                        |   | 011.01101.110    | HIIIMACCC       |
| 251 | CGTAACCGAG                              | TAAGATTTGG             | CCATTTTCGC                              | GGGAAAACTG       | AATAAGAGGA      |
|     | GCATTGGCTC                              | ATTCTAAACC             | GGTAAAAGCG                              | CCCTTTTGAC       | TTATTCTCCT      |
|     |   |                        |   |                  |                 |
| 301 | AGTGAAATCT                              | GAATAATTTT             | GTGTTACTCA                              | TAGCGCGTAA       | TATTTGTCTA      |
|     | TCACTTTAGA                              | CTTATTAAAA             | CACAATGAGT                              | ATCGCGCATT       | ATAAACAGAT      |
|     |   | •                      |   |                  |                 |
| 351 |   |                        |   | AGACTCGCCC       |                 |
|     | CCCGGCGCCC                              | CTGAAACTGG             | CAAATGCACC                              | TCTGAGCGGG       | TCCACAAAAA      |
| 401 | OPG LOCKOPP                             | PP00000000000          |   |                  |                 |
| 401 |   |                        |   | TTGGCGTTTT       |                 |
| -   | GAGICCACAA                              | MAGGEGEAAG             | GCCCAGTTTC                              | AACCGCAAAA       | TAATAATATC      |
| 451 | פרבפררפרפ                               | ጥርር ልጥጥርር ልጥ           | <b>እ</b> ርርጥጥር ጥአጥር                     | CATATCATAA       | TATCTACATT      |
|     |   |                        |   | GTATAGTATT       |                 |
|     | *************************************** |                        | TOGETICATAG                             | GININGINII       | MINCHIGIAN      |
| 501 | TATATTGGCT                              | CATGTCCAAC             | ATTACCGCCA                              | TGTTGACATT       | GATTATTGAC      |
|     |   |                        |   | ACAACTGTAA       |                 |
|     |   |                        |   |                  |                 |
| 551 | TACTTATTAA                              | TAGTAATCAA             | TTACGGGGTC                              | ATTAGTTCAT       | AGCCCATATA      |
|     | ATCAATAATT                              | ATCATTAGTT             | AATGCCCCAG                              | TAATCAAGTA       | TCGGGTATAT      |
|     |   | •                      |   |                  |                 |
| 601 |   |                        |   | ATGGCCCGCC       |                 |
|     | ACCTCAAGGC                              | GCAATGTATT             | GAATGCCATT                              | TACCGGGCGG       | ACCGACTGGC      |
|     |   |                        |   |                  |                 |
| 651 |   |                        |   | ATGACGTATG       |                 |
|     | GGGTTGCTGG                              | GGGCGGGTAA             | CTGCAGTTAT                              | TACTGCATAC       | AAGGGTATCA      |
| 701 | AACGCCAATA                              | CCCACTOTOCC            | 3 mmc 3 ccmc 3                          | **************** | m > mmm > 0.00m |
|     | TTGCGGTTAT                              |                        |   |                  |                 |
| •   | 1100001171                              | cccidanao              | IAAC1GCAG1                              | TACCCACC IC      | AIAAAIGCCA      |
| 751 | AAACTGCCCA                              | CTTGGCAGTA             | CATCAAGTGT                              | ATCATATICC       | AAGTACGCCC      |
|     |   |                        |   | TAGTATACGG       |                 |
|     |   |                        |   |                  |                 |
| 801 | CCTATTGACG                              | TCAATGACGG             | TAAATGGCCC                              | GCCTGGCATT       | ATGCCCAGTA      |
|     | GGATAACTGC                              | AGTTACTGCC             | ATTTACCGGG                              | CGGACCGTAA       | TACGGGTCAT      |
|     |   |                        |   |                  |                 |
| 851 | CATGACCTTA                              |                        |   |                  |                 |
|     | GTACTGGAAT                              | ACCCTGAAAG             | GATGAACCGT                              | CATGTAGATG       | CATAATCAGT      |

7 i jure 26A

| 901  | TCGCTATTAC<br>AGCGATAATG | GGGGATG<br>GTACCACTAC    | CGGTTTTGGC<br>GCCAAAACCG | AGTACATCAA<br>TCATGTAGTT     | TGGGCC LA<br>ACCCGCACCT      |
|------|--------------------------|--------------------------|--------------------------|------------------------------|------------------------------|
| 951  | TAGCGGTTTG<br>ATCGCCAAAC | ACTCACGGGG<br>TGAGTGCCCC | ATTTCCAAGT<br>TAAAGGTTCA | CTCCACCCCA<br>GAGGTGGGGT     | TTGACGTCAA<br>AACTGCAGTT     |
| 1001 | TGGGAGTTTG<br>ACCCTCAAAC | TTTTGGCACC<br>AAAACCGTGG | AAAATCAACG<br>TTTTAGTTGC | GGACTTTCCA<br>CCTGAAAGGT     | AAATGTCGTA<br>TTTACAGCAT     |
| 1051 | ACAACTCCGC<br>TGTTGAGGCG | CCCATTGACG<br>GGGTAACTGC | CAAATGGGCG<br>GTTTACCCGC | GTAGGCGTGT<br>CATCCGCACA     | ACGGTGGGAG<br>TGCCACCCTC     |
| 1101 | CAGATATATT               | CGTCTCGAGC               | AAATCACTTG               | CGTCAGATCG<br>GCAGTCTAGC     | GGACCTCTGC                   |
| 1151 | GGTAGGTGCG               | ACAAAACTGG               | AGGTATCTTC               | ACACCGGGAC<br>TGTGGCCCTG     | GCTAGGTCGG                   |
| 1201 | AGGCGCCGGC               | CCTTGCCACG               | TAACCTTGCG               | GGATTCCCCG<br>CCTAAGGGGC     | ACGGTTCTCA                   |
| 1251 | CTCTAGATGG               | TACCGGGGGT               | AGAGGGGGTA               | TGAGACTGTG<br>ACTCTGACAC     | GGACACTTCG                   |
| 1301 | ACTTCGGACC               | GTACCTACCG               | GGGTTCCACT               | AGCAGTGGCC<br>TCGTCACCGG     | GGACTGACTC                   |
| 1351 | CTCTTCTAGT               | TCCGGGACCA               | CCTTTAGACG               | ACTGAGATGG<br>TGACTCTACC     | TCTTCCTCCC                   |
| 1401 | GTTTTAGAGG               | TTCTAACCGG               | GGCTCTTGGG               | CTACAACACC<br>GATGTTGTGG     | GGACACAAAC                   |
| 1451 | GGTAGTTCTT               | CTTCCTGAGG               | TGGTTCACCT               | GGAAGCTGGT                   | CCTGAAGTCC                   |
| 1501 | CTCGACTTGT               | TCTCCTGGGI               | CCTGAAGACC               | CTCCACGTCG                   |                              |
| 1551 | GGTGGGGCGA               | CCGGACTTCI               | TCTTCTTCAG               | ACACTGACAC                   | CTGGCTGTGG                   |
| 1601 | CCCTACGGAT               | GAAGAGACAC               | GGGGACCTAC               | : TCCTGAAGIC                 | GAAGTACACT<br>CTTCATGTGA     |
| 1651 | CGGAAGTGG1               | AGGGGAGGT                | GTTGTTACT(               | TGGGGACCGT                   | TCAGGTACCA AGTCCATGGT        |
|      | CATGTTACAC               | GACGGGGTCC               | CGACCTTCCC               | GAGGGGACGG                   | : ATCTTCCAGT<br>: TAGAAGGTCA |
|      | GGAGGTACT                | GTTCTAGGAC               | CTCGGGAAG                | CCTTCGTCT1                   | CCCTGACATT<br>GGGACTGTAA     |
| 1801 | GTGATCTAC(<br>CACTAGATG( | AGTACATGG<br>CATGTACC    | TGCCCTGTAT               | r GTGGGCTCTC<br>A CACCCGAGAC | ACCTGGAGAT<br>TGGACCTCTA     |



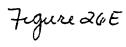
| 1851 |                          | A CCAAGA<br>TCCTGGTTCT   |      |                          |
|------|--------------------------|--------------------------|------|--------------------------|
| 1901 |                          | CACCCCTGAC<br>GTGGGGACTG | <br> |                          |
| 1951 |                          | ATGAGCTGCA<br>TACTCGACGT | <br> |                          |
| 2001 |                          | AAGGACTCCT<br>TTCCTGAGGA |      |                          |
| 2051 |                          | CTGGGCCTCC<br>GACCCGGAGG |      |                          |
| 2101 |                          | TGCTGAGGGG<br>ACGACTCCCC |      |                          |
| 2151 |                          | GCTGAGCTGG<br>CGACTCGACC | <br> |                          |
| 2201 |                          | TGGGGTGTAC<br>ACCCCACATG | <br> |                          |
| 2251 |                          | AGGGCCAGGG<br>TCCCGGTCCC | <br> |                          |
| 2301 |                          | CTGAAGACTG<br>GACTTCTGAC |      |                          |
| 2351 |                          | GAAGCAGCTG<br>CTTCGTCGAC | <br> |                          |
| 2401 | TCCATTGTGA<br>AGGTAACACT | TCTGGGGCAA<br>AGACCCCGTT | <br> |                          |
| 2451 |                          | GAGACCTGGT<br>CTCTGGACCA |      |                          |
| 2501 |                          | GTTTGTGAAC<br>CAAACACTTG |      |                          |
| 2551 |                          | AGCCCATTGT<br>TCGGGTAACA | <br> |                          |
| 2601 | TGCCAACAGG<br>ACGGTTGTCC |                          |      | ACCAACAGGG<br>TGGTTGTCCC |
| 2651 | GCAGGCAGAA<br>CGTCCGTCTT |                          |      | GAAGACTGCC<br>CTTCTGACGG |
| 2701 | CTCCAGGCCA<br>GAGGTCCGGT |                          |      | AGGTGAACAT<br>TCCACTTGTA |
| 2751 | TGTGACTGCC<br>ACACTGACGG |                          |      | CAGCCTGATC<br>GTCGGACTAG |



| 2801 | AGTCTGAGTC | T CTGGTG                 | AACCAGATCA | TTGAGCAGCT | GATCAA G                 |
|------|------------|--------------------------|------------|------------|--------------------------|
|      | TCAGACTCAG | ACTCGACCAC               | TTGGTCTAGT | AACTCGTCGA | CTAGTTCTTC               |
| 2851 | GAGAAGGTGT | ACCTGGCCTG               | GGTGCCTGCC | CACAAGGGCA | TTGGGGGCAA               |
|      | CTCTTCCACA | TGGACCGGAC               | CCACGGACGG | GTGTTCCCGT | AACCCCCGTT               |
| 2901 |            | GACAAGCTGG<br>CTGTTCGACC |            |            |                          |
| 2951 |            | TGACAAGGCC<br>ACTGTTCCGG |            |            | CCACTCCAAC GGTGAGGTTG    |
| 3001 | TGGAGGGCTA | TGGCCTCTGA               | CTTCAACCTG | CCCCTGTGG  | TGGCTAAGGA               |
|      | ACCTCCCGAT | ACCGGAGACT               | GAAGTTGGAC | GGGGGACACC | ACCGATTCCT               |
| 3051 |            | TCCTGTGACA<br>AGGACACTGT |            |            |                          |
| 3101 |            | CTGCTCCCCT<br>GACGAGGGGA |            |            | CACCCACCTG<br>GTGGGTGGAC |
| 3151 | GAGGGCAAGG | TGATCCTGGT               | GGCTGTGCAT | GTGGCCTCCG | GCTACATTGA               |
|      | CTCCCGTTCC | ACTAGGACCA               | CCGACACGTA | CACCGGAGGC | CGATGTAACT               |
| 3201 |            | ATCCCTGCTG<br>TAGGGACGAC |            |            |                          |
| 3251 | TGAAGCTGGC | TGGCAGGTGG               | CCTGTGAAGA | CCATCCACAC | TGCCAATGGC               |
|      | ACTTCGACCG | ACCGTCCACC               | GGACACTTCT | GGTAGGTGTG | ACGGTTACCG               |
| 3301 |            |                          |            |            | GGGCTGGCAT<br>CCCGACCGTA |
| 3351 |            | TTTGGCATCC<br>AAACCGTAGG |            |            | GGGGTGGTGG<br>CCCCACCACC |
| 3401 |            |                          |            |            | GAGGGACCAG<br>CTCCCTGGTC |
| 3451 |            |                          |            |            | TCCACAACTT<br>AGGTGTTGAA |
| 3501 | CAAGAGGAAG | GGGGGCATCG               | GGGGCTACTC | CGCTGGGGAG | AGGATTGTGG               |
|      | GTTCTCCTTC | CCCCCGTAGC               | CCCCGATGAG | GCGACCCCTC | TCCTAACACC               |
| 3551 | ACATCATTGC | CACAGACATC               | CAGACCAAGG | AGCTCCAGAA | GCAGATCACC               |
|      | TGTAGTAACG | GTGTCTGTAG               | GTCTGGTTCC | TCGAGGTCTT | CGTCTAGTGG               |
| 3601 | AAGATCCAGA | ACTTCAGGGT               | GTACTACAGG | GACTCCAGGA | ACCCCCTGTG               |
|      | TTCTAGGTCT | TGAAGTCCCA               | CATGATGTCC | CTGAGGTCCT | TGGGGGACAC               |
| 3651 | GAAGGGCCCT | GCCAAGCTGC               | TGTGGAAGGG | GGAGGGGCT  | GTGGTGATCC               |
|      | CTTCCCGGGA | CGGTTCGACG               | ACACCTTCCC | CCTCCCCGA  | CACCACTAGG               |
| 3701 | AGGACAACTC | TGACATCAAG               | GTGGTGCCCA | GGAGGAAGGC | CAAGATCATC               |
|      | TCCTGTTGAG | ACTGTAGTTC               | CACCACGGGT | CCTCCTTCCG | GTTCTAGTAG               |

Figure 26 D

| 3751 | AGGGACTATG<br>TCCCTGATAC | AGCAGAT<br>CO.TCGTCTA    | GGCTGGGGAT<br>CCGACCCCTA | GACTGTGTGG<br>CTGACACACC | CCTCCA CCA<br>GGAGGT GT    |
|------|--------------------------|--------------------------|--------------------------|--------------------------|----------------------------|
| 3801 |                          |                          |                          | TGTGCCTTCT<br>ACACGGAAGA |                            |
| 3851 | CATCTGTTGT<br>GTAGACAACA |                          |                          | CCTTGACCCT<br>GGAACTGGGA |                            |
| 3901 |                          |                          |                          | GAAATTGCAT<br>CTTTAACGTA |                            |
| 3951 | GAGTAGGTGT<br>CTCATCCACA |                          |                          | GGTGGGGCAG<br>CCACCCGTC  |                            |
| 4001 | •••••                    |                          |                          | CTGGGGATGC<br>GACCCCTACG |                            |
| 4051 |                          |                          |                          | GTGGGCGTGG<br>CACCCGCACC |                            |
| 4101 |                          |                          |                          | TAGTTTTGTA<br>ATCAAAACAT |                            |
| 4151 |                          |                          |                          | GTTTGATGGA<br>CAAACTACCT |                            |
| 4201 | GCTCATATTT<br>CGAGTATAAA |                          |                          | GGGCCGGGGT<br>CCCGGCCCCA |                            |
| 4251 |                          |                          |                          | GTCCTGCCCG<br>CAGGACGGGC | CAAACTCTAC<br>GTTTGAGATG . |
| 4301 |                          |                          |                          | GCCGTTGGAG<br>CGGCAACCTC |                            |
| 4351 | CCCCCCCCCC               | TTCAGCCGCT<br>AAGTCGGCGA | GCAGCCACCG<br>CGTCGGTGGC | CCCGCGGGAT<br>GGGCGCCCTA | TGTGACTGAC<br>ACACTGACTG   |
| 4401 |                          |                          |                          | GCAGCTTCCC<br>CGTCGAAGGG |                            |
| 4451 | CCGCGATGAC<br>GGCGCTACTG | AAGTTGACGG<br>TTCAACTGCC | CTCTTTTGGC<br>GAGAAAACCG | ACAATTGGAT<br>TGTTAACCTA | TCTTTGACCC<br>AGAAACTGGG   |
| 4501 |                          |                          |                          |                          | CCAGCAGGTT<br>GGTCGTCCAA   |
| 4551 | TCTGCCCTGA<br>AGACGGGACT | AGGCTTCCTC<br>TCCGAAGGAG | CCCTCCCAAT<br>GGGAGGGTTA | GCGGTTTAAA<br>CGCCAAATTT | ACATAAATAA<br>TGTATTTATT   |
| 4601 |                          |                          |                          |                          | TGCTGTCTTT<br>ACGACAGAAA   |
|      |                          |                          |                          |                          | GTCTCGGTCG<br>CAGAGCCAGC   |



| 4701 | TTGAGGGTCC<br>AACTCCCAGG | TCTCTATTTT<br>ATAAAA     | TTCCAGGACG<br>AAGGTCCTGC | TGGTAAAGGT<br>ACCATTTCCA | CTGAGA A                 |
|------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| 4751 |                          |                          |                          | GGGGTGGAGG<br>CCCCACCTCC |                          |
| 4801 |                          |                          |                          | AGATGATCCA<br>TCTACTAGGT |                          |
| 4851 |                          |                          |                          | TTCAGTAGCA<br>AAGTCATCGT |                          |
| 4901 |                          |                          |                          | AAAGCGGTTA<br>TTTCGCCAAT |                          |
| 4951 |                          |                          |                          | TGGACTGTAT<br>ACCTGACATA |                          |
| 5001 |                          |                          |                          | TTCATGTTGT<br>AAGTACAACA |                          |
| 5051 |                          |                          |                          | TTTGTCATGT<br>AAACAGTACA |                          |
| 5101 |                          |                          |                          | TGTGACCTCC<br>ACACTGGAGG |                          |
| 5151 |                          |                          |                          | CCACGGGCGG<br>GGTGCCCGCC |                          |
| 5201 |                          |                          |                          | GTTGTGTTCC<br>CAACACAAGG |                          |
| 5251 |                          |                          |                          | GGAGGGTGCC<br>CCTCCCACGG |                          |
| 5301 |                          |                          |                          | TTACCCTCAC<br>AATGGGAGTG |                          |
| 5351 |                          |                          |                          | CATGTCTACC<br>GTACAGATGG |                          |
| 5401 | TGAAGAAAAC<br>ACTTCTTTTG | GGTTTCCGGG<br>CCAAAGGCCC | GTAGGGGAGA<br>CATCCCCTCT | TCAGCTGGGA<br>AGTCGACCCT | AGAAAGCAGG<br>TCTTTCGTCC |
| 5451 | TTCCTGAGCA<br>AAGGACTCGT | GCTGCGACTT<br>CGACGCTGAA | ACCGCAGCCG<br>TGGCGTCGGC | GTGGGCCCGT<br>CACCCGGGCA | AAATCACACC<br>TTTAGTGTGG |
| 5501 | TATTACCGGC<br>ATAATGGCCG | TGCAACTGGT<br>ACGTTGACCA | AGTTAAGAGA<br>TCAATTCTCT | GCTGCAGCTG<br>CGACGTCGAC | CCGTCATCCC<br>GGCAGTAGGG |
| 5551 | TGAGCAGGGG<br>ACTCGTCCCC | GGCCACTTCG<br>CCGGTGAAGC | TTAAGCATGT<br>AATTCGTACA | CCCTGACTCG<br>GGGACTGAGC | CATGTTTTCC<br>GTACAAAAGG |
| 5601 | CTGACCAAAT<br>GACTGGTTTA | CCGCCAGAAG<br>GGCGGTCTTC | GCGAGCGGC                | CCCAGCGATA<br>GGGTCGCTAT | GCAGTTCTTG<br>CGTCAAGAAC |

Figure 26F

| 5651 | CAAGGAAGCA<br>GTTCCTTCGT |   | ACCGTCCGCC<br>TGGCAGGCGG     |                          |
|------|--------------------------|---|------------------------------|--------------------------|
| 5701 |                          |   | GGTCCCACAG<br>CCAGGGTGTC     |                          |
| 5751 |                          |   | <br>CCTCGTTTCG<br>GGAGCAAAGC |                          |
| 5801 |                          |   | <br>TCGTCCAGAC<br>AGCAGGTCTG |                          |
|      | CATGTCTTTC<br>GTACAGAAAG |   |                              |                          |
| 5901 |                          | _ | CCAGGGTGCG<br>GGTCCCACGC     |                          |
| 5951 |                          |   | <br>TCGCCCTGCG<br>AGCGGGACGC |                          |
| 6001 |                          |   | CCCCTCCGCG<br>GGGGAGGCGC     |                          |
| 6051 |                          |   | CGCACGAGGG<br>GCGTGCTCCC     |                          |
| 6101 |                          |   | AATACCGATT<br>TTATGGCTAA     |                          |
| 6151 |                          |   | <br>CTCGCATTCC<br>GAGCGTAAGG |                          |
| 6201 | •                        |   | GGTTTCCCCC<br>CCAAAGGGGG     |                          |
| 6251 |                          |   | <br>CGGTGTCCAC<br>GCCACAGGTG |                          |
| 6301 |                          |   | <br>CTTGAGAGGC<br>GAACTCTCCG |                          |
| 6351 |                          |   | <br>ACTCGGACCA<br>TGAGCCTGGT |                          |
| 6401 | AAGGCTCGCG<br>TTCCGAGCGC |   | GCTAAGTGGG<br>CGATTCACCC     |                          |
| 6451 | GTCGTTGTCC<br>CAGCAACAGG |   |                              | AGACACATGT<br>TCTGTGTACA |
| 6501 | CGCCCTCTTC<br>GCGGGAGAAG |   |                              | GTAGGCCACG<br>CATCCGGTGC |
| 6551 | TGACCGGGTG<br>ACTGGCCCAC |   |                              | GGGCGCGTTC<br>CCCGCGCAAG |

Figure 266

| 6601  | GTCCTCACTC<br>CAGGAGTGAG | TCTTCCGCAT<br>AGGCGTA    | CGCTGTCTGC<br>GCGACAGACG | CAGGGCCAGG<br>CTCCCGGTCG     | ACAACO C                 |
|-------|--------------------------|--------------------------|--------------------------|------------------------------|--------------------------|
| 6651  | AGTACTCCCT               | CTGAAAAGCG               | GGCATGACTT               | CTGCGCTAAG                   | ATTGTCAGTT               |
|       | TCATGAGGGA               | GACTTTTCGC               | CCGTACTGAA               | GACGCGATTC                   | TAACAGTCAA               |
| 6701  | TCCAAAAACG               | AGGAGGATTT               | GATATTCACC               | TGGCCCGCGG                   | TGATGCCTTT               |
|       | AGGTTTTTGC               | TCCTCCTAAA               | CTATAAGTGG               | ACCGGGCGCC                   | ACTACGGAAA               |
| 6751  | GAGGGTGGCC               | GCATCCATCT               | GGTCAGAAAA               | GACAATCTTT                   | TTGTTGTCAA               |
|       | CTCCCACCGG               | CGTAGGTAGA               | CCAGTCTTTT               | CTGTTAGAAA                   | AACAACAGTT               |
| 6801  | GCTTGGTGGC               | AAACGACCCG               | TAGAGGGCGT               | TGGACAGCAA                   | CTTGGCGATG               |
|       | CGAACCACCG               | TTTGCTGGGC               | ATCTCCCGCA               | ACCTGTCGTT                   | GAACCGCTAC               |
| 6851  | GAGCGCAGGG               | TTTGGTTTTT               | GTCGCGATCG               | GCGCGCTCCT                   | TGGCCGCGAT               |
|       | CTCGCGTCCC               | AAACCAAAAA               | CAGCGCTAGC               | CGCGCGAGGA                   | ACCGGCGCTA               |
| 6901  | GTTTAGCTGC               | ACGTATTCGC               | GCGCAACGCA               | CCGCCATTCG                   | GGAAAGACGG               |
|       | CAAATCGACG               | TGCATAAGCG               | CGCGTTGCGT               | GGCGGTAAGC                   | CCTTTCTGCC               |
| 6951  | TGGTGCGCTC               | GTCGGGCACC               | AGGTGCACGC               | GCCAACCGCG                   | GTTGTGCAGG               |
|       | ACCACGCGAG               | CAGCCCGTGG               | TCCACGTGCG               | CGGTTGGCGC                   | CAACACGTCC               |
| 7001  | GTGACAAGGT               | CAACGCTGGT               | GGCTACCTCT               | CCGCGTAGGC                   | GCTCGTTGGT               |
|       | CACTGTTCCA               | GTTGCGACCA               | CCGATGGAGA               | GGCGCATCCG                   | CGAGCAACCA               |
| .7051 | CCAGCAGAGG               | CGGCCGCCCT               | TGCGCGAGCA               | GAATGGCGGT                   | AGGGGGTCTA               |
|       | GGTCGTCTCC               | GCCGGCGGGA               | ACGCGCTCGT               | CTTACCGCCA                   | TCCCCCAGAT               |
| 7101  | GCTGCGTCTC               | GTCCGGGGGG               | TCTGCGTCCA               | CGGTAAAGAC                   | CCCGGGCAGC               |
|       | CGACGCAGAG               | CAGGCCCCCC               | AGACGCAGGT               | GCCATTTCTG                   | GGGCCCGTCG               |
| 7151  | AGGCGCGCGT               | CGAAGTAGTC               | TATCTTGCAT               | CCTTGCAAGT                   | CTAGCGCCTG               |
|       | TCCGCGCGCA               | GCTTCATCAG               | ATAGAACGTA               | GGAACGTTCA                   | GATCGCGGAC               |
| 7201  | CTGCCATGCG<br>GACGGTACGC | CGGGCGGCAA<br>GCCCGCCGTT | GCGCGCGCGAG              | GTATGGGTTG<br>CATACCCAAC     | AGTGGGGGAC<br>TCACCCCTG  |
| 7251  | CCCATGGCAT<br>GGGTACCGTA | CCCACCCAC                | AGCGCGGAGG<br>TCGCGCCTCC | CGTACATGCC<br>CGCATGTACGG    | GCAAATGTCG<br>CGTTTACAGC |
| 7301  | TAAACGTAGA<br>ATTTGCATCT | GGGGCTCTCT               | GAGTATTCCA<br>CTCATAAGGT | AGATATGTAG<br>OTADATATOT     | GGTAGCATCT<br>CCATCGTAGA |
| 7351  | TCCACCGCGG<br>AGGTGGCGCC | ATGCTGGCGC<br>TACGACCGCG | GCACGTAAT(               | GTATAGTTCG<br>CATATCAAGC     | TGCGAGGGAG<br>ACGCTCCCTC |
| 7401  | CGAGGAGGTC<br>GCTCCTCCAG | GGGACCGAGG               | TTGCTACGGG               | CGGGCTGCTC                   | TGCTCGGAAG ACGAGCCTTC    |
| 7451  | ACTATCTGCC<br>TGATAGACGG | TGAAGATGGC               | ATGTGAGTTY<br>TACACTCAA  | GATGATATGO<br>CTACTATACO     | TTGGACGCTG<br>AACCTGCGAC |
| 7501  | GAAGACGTTG<br>CTTCTGCAAC | AAGCTGGCGI               | CTGTGAGACO               | C TACCGCGTCA<br>G ATGGCGCAGT | CGCACGAAGG<br>CGTGCTTCC  |

Figure 26 H

| 7551 | AGGCGTAGGA<br>TCCGCATCCT | GCGCAGC<br>CAGCGCGTCG    |     |                          |
|------|--------------------------|--------------------------|-----|--------------------------|
| 7601 |                          | AGTAGTCCAG<br>TCATCAGGTC |     |                          |
| 7651 |                          | TTCCACAGCT<br>AAGGTGTCGA |     |                          |
| 7701 |                          | TTGGATCGGA<br>AACCTAGCCT |     |                          |
| 7751 |                          | ACTGGTTGAC<br>TGACCAACTG |     |                          |
| 7801 |                          | TATGCCTGCG<br>ATACGGACGC |     | <br>                     |
| 7851 |                          | CCTGACCATG<br>GGACTGGTAC |     |                          |
| 7901 |                          | CGCCCTGCTC<br>GCGGGACGAG |     |                          |
| 7951 |                          | GGCAGGGCGA<br>CCGTCCCGCT |     |                          |
| 8001 |                          | AAAGTTGCGT<br>TTTCAACGCA | +-, | <br>                     |
| 8051 |                          | TTACCTGGGC<br>ÄATGGACCCG |     |                          |
| 8101 |                          | ACAATGTAAA<br>TGTTACATTT |     |                          |
| 8151 |                          | TTTAAGTTCC<br>AAATTCAAGG |     |                          |
| 8201 |                          | AAAGGGCCCA<br>TTTCCCGGGT |     |                          |
| 8251 |                          | AGGTCACGGG<br>TCCAGTGCCC |     | <br>                     |
| 8301 |                          |                          |     | GCAGTAGAAG<br>CGTCATCTTC |
| 8351 |                          |                          |     | CGGCTAGGTC<br>GCCGATCCAG |
| 8401 |                          |                          |     | ATGACCAGCA<br>TACTGGTCGT |
| 8451 |                          |                          |     | ATAGGTCTCT<br>TATCCAGAGA |

Figure 26I

| 8501 | ACATCGTAGG<br>TGTAGCATCC | TAAAGAG<br>ACTGTTTCTC     | ACGCTCGGTG<br>TGCGAGCCAC                | CGAGGATGCG<br>GCTCCTACGC     | AGCCGA GGG<br>TCGGCTAGCC |
|------|--------------------------|---------------------------|---|------------------------------|--------------------------|
| 8551 | GAAGAACTGG<br>CTTCTTGACC | ATCTCCCGCC<br>TAGAGGGCGG  | ACCAATTGGA<br>TGGTTAACCT                | GGAGTGGCTA<br>CCTCACCGAT     | TTGATGTGGT<br>AACTACACCA |
| 8601 | CTTTCATCTT               | CAGGGACGCT                | GCCCGCCTTG                              | ACTCGTGCTG<br>TGAGCACGAC     | CGAAAACATT               |
| 8651 | TTTGCACGCG               | TCATGACCGT                | CGCCACGTGC                              | GGCTGTACAT<br>CCGACATGTA     | GGACGTGCTC               |
| 8701 | GTTGACCTGA<br>CAACTGGACT | CGACCGCGCA<br>GCTGGCGCGT  | CAAGGAAGCA<br>GTTCCTTCGT                | GAGTGGGAAT<br>CTCACCCTTA     | TTGAGCCCCT<br>AACTCGGGGA |
| 8751 | CGCCTGGCGG<br>GCGGACCGCC | GTTTGGCTGG<br>CAAACCGACC  | TGGTCTTCTA<br>ACCAGAAGAT                | CTTCGGCTGC<br>GAAGCCGACG     | TTGTCCTTGA<br>AACAGGAACT |
| 8801 | CCGTCTGGCT<br>GGCAGACCGA | GCTCGAGGGG<br>CGAGCTCCCC  | AGTTACGGTG<br>TCAATGCCAC                | GATCGGACCA<br>CTAGCCTGGT     | CCACGCCGCG<br>GGTGCGGCGC |
| 8851 | CGAGCCCAAA<br>GCTCGGGTTT | GTCCAGATGT<br>CAGGTCTACA  | CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC | CGGTCGGAGC<br>GCCAGCCTCG     | TTGATGACAA<br>AACTACTGTT |
| 8901 | CATCGCGCAG<br>GTAGCGCGTC | ATGGGAGCTG<br>TACCCTCGAC  | TCCATGGTCT<br>AGGTACCAGA                | GGAGCTCCCG<br>CCTCGAGGGC     | CGGCGTCAGG<br>GCCGCAGTCC |
| 8951 | TCAGGCGGGA<br>AGTCCGCCCT | GCTCCTGCAG<br>CGAGGACGTC  | GTTTACCTCG<br>CAAATGGAGC                | CATAGACGGG<br>GTATCTGCCC     | TCAGGGCGCG<br>AGTCCCGCGC |
| 9001 | GGCTAGATCC<br>CCGATCTAGG | AGGTGATACC<br>TCCACTATGG  | TAATTTCCAG<br>ATTAAAGGTC                | GGGCTGGTTG                   | GTGGCGGCGT<br>CACCGCCGCA |
| 9051 | CGATGGCTTG<br>GCTACCGAAC | CAAGAGGCCG<br>GTTCTCCGGC  | CATCCCCGCG<br>GTAGGGGCGC                | GCGCGACTAC<br>CGCGCTGATG     | GGTACCGCGC<br>CCATGGCGCG |
| 9101 | CCGCCCGCCA               | ccceccccc                 | GGTGTCCTTG                              | GATGATGCAT<br>CTACTACGTA     | CTAAAAGCGG<br>GATTTTCGCC |
| 9151 | TGACGCGGGC<br>ACTGCGCCCG | GAGCCCCCGG<br>CTCGGGGGCC  | AGGTAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG  | GGCTCCGGAC<br>CCGAGGCCTG     | CCGCCGGGAG<br>GGCGGCCCTC |
| 9201 | AGGGGGCAGG<br>TCCCCGTCC  | GGCACGTCGG<br>CCGTGCAGCC  | cecececece                              | GGGCAGGAGC<br>CCCGTCCTCG     | TGGTGCTGCG<br>ACCACGACGC |
| 9251 | CGCGTAGGTT               | GCTGGCGAAC                | GCGACGACGC<br>GCTGCTGCC                 | GGCGGTTGAT<br>CCGCCAACTA     | CTCCTGAATC<br>GAGGACTTAG |
| 9301 | TGGCGCCTCT<br>ACCGCGGAGA | CGCACTTCTC                | CTGCCCGGG                               | GTGAGCTTGA<br>CACTCGAACT     | ACCTGAAAGA<br>TGGACTTTCT |
| 9351 | GAGTTCGACA<br>CTCAAGCTGT | A GAATCAATTI<br>CTTAGTTAA | CGGTGTCGT:                              | T GACGGCGCCGG                | TGGCGCAAAA<br>ACCGCGTTTT |
| 9401 | TCTCCTGCA(<br>AGAGGACGT( | CAGAGGACTO                | TTGTCTTGAT                              | r aggcgatcto<br>A tccgctagao | GGCCATGAAC<br>CCGGTACTTG |

Figure 26 J

| 9451  | TGCTCGATCT               |                          |              |              |                          |
|-------|--------------------------|--------------------------|--------------|--------------|--------------------------|
|       | ACGAGCTAGA               | GAAGGAGGAC               | CTCTAGAGGC   | GCAGGCCGAG   | CGAGGTGCCA               |
| 9501  |                          | TCGTTGGAAA               |              |              |                          |
|       | CCGCCGCTCC               | AGCAACCTTT               | ACGCCCGGTA   | CTCGACGCTC   | TTCCGCAACT               |
| 9551  | GGCCTCCCTC               |                          |              |              |                          |
|       | CCGGAGGGAG               | CAAGGTCTGC               | GCCGACATCT   | GGTGCGGGG    | AAGCCGTAGC               |
| 9601  |                          | TGACCACCTG<br>ACTGGTGGAC |              |              |                          |
|       | GCCCGCGT                 | ACTGGTGGAC               | GUGUTUTAAU   | TUGAGGTGCA   | CGGCCCGCTT               |
| 9651  | GACGGCGTAG               |                          |              |              |                          |
|       | CTGCCGCATC               | AAAGCGTCCG               | CGACTTTCTC   | CATCAACTCC   | CACCACCGCC               |
| 9701  | TGTGTTCTGC               | CACGAAGAAG               | TACATAACCC   | AGCGTCGCAA   | CGTGGATTCG               |
|       | ACACAAGACG               | GTGCTTCTTC               | ATGTATTGGG   | TCGCAGCGTT   | GCACCTAAGC               |
| 9751  | TTGATATCCC               | CCAAGGCCTC               | AAGGCGCTCC   | ATGGCCTCGT   | AGAAGTCCAC               |
|       | AACTATAGGG               | GGTTCCGGAG               | TTCCGCGAGG   | TACCGGAGCA   | TCTTCAGGTG               |
| 9801  | GGCGAAGTTG               | AAAAACTGGG               | AGTTGCGCGC   | CGACACGGTT   | AACTCCTCCT               |
|       | CCGCTTCAAC               | TTTTTGACCC               | TCAACGCGCG   | GCTGTGCCAA   | TTGAGGAGGA               |
| 9851  | CCAGAAGACG               | GATGAGCTCG               | GCGACAGTGT   | CGCGCACCTC   | GCGCTCAAAG               |
|       | GGTCTTCTGC               | CTACTCGAGC               | CGCTGTCACA   | GCGCGTGGAG   | CGCGAGTTTC               |
| 9901  | GCTACAGGGG               | CCTCTTCTTC               | TTCTTCAATC   | TCCTCTTCCA   | TAAGGGCCTC               |
|       | CGATGTCCCC               | GGAGAAGAAG               | AAGAAGTTAG   | AGGAGAAGGT   | ATTCCCGGAG               |
| 9951  | CCCTTCTTCT               | TCTTCTGGCG               | GCGGTGGGGG   | AGGGGGGACA   | CGGCGGCGAC               |
|       | GGGAAGAAGA               | AGAAGACCGC               | CGCCACCCCC   | TCCCCCCTGT   | GCCGCCGCTG               |
| 10001 | GACGGCGCAC               | CGGGAGGCGG               | TCGACAAAGC   | GCTCGATCAT   | CTCCCCGCGG               |
|       | CTGCCGCGTG               | GCCCTCCGCC               | AGCTGTTTCG   | CGAGCTAGTA   | GAGGGGCGCC               |
| 10051 | CGACGGCGCA               | TGGTCTCGGT               | GACGGCGCGG   | CCGTTCTCGC   | GGGGGCGCAG               |
|       | GCTGCCGCGT               | ACCAGAGCCA               | CTGCCGCGCC   | GGCAAGAGCG   | CCCCCCCCTC               |
| 10101 | TTEGAAGACG               | CCGCCCGTCA               | TGTCCCGGTT   | ATGGGTTGGC   | GGGGGCTGC                |
|       |                          | GGCGGGCAGT               |              |              |                          |
| 10151 | ראיזיכרככראכ             | CCATACCCC                | CTAACCATCC   | እጥሮጥሮአአሮአአ   | TTGTTGTGTA               |
| 10151 |                          | CCTATGCCGC               |              |              |                          |
| 10201 | CCDXCDCCCC               | CCCCCACCCA               | CCTCXCCCXC   | TO CO CATOCA | CCGGATCGGA               |
| 10201 |                          |                          |              |              | GGCCTAGCCT               |
| 10051 | 1110000000               | 3.0333.00000             | CD2 2CC2 CDC | NCNC00000X   | COMP COOMC 3             |
| 10251 | AAACCTCTCG<br>TTTGGAGAGC |                          |              |              | CCATCCGACT               |
|       | 001000000                | 0000000000               | 000000000    | ACAAA        | Benedera e               |
|       | CGTGGCACCG               |                          |              |              | TCTGGCGGAG<br>AGACCGCCTC |
|       |                          |                          |              |              |                          |
| 10351 |                          |                          |              |              | GGCGGATGGT<br>CCGCCTACCA |
|       |                          |                          |              |              |                          |

Figure 26 K

| 10401 | CGACAGAAGC<br>GCTGTCTTCG | A TGTCCT<br>TACAGGA       | TGGGTCCGGC<br>ACCCAGGCCG | CTGCTGAATG<br>GACGACTTAC  | CGCAGGGTT<br>CCCCCCA     |
|-------|--------------------------|---------------------------|--------------------------|---------------------------|--------------------------|
| 10451 | CGGCCATGCC<br>GCCGGTACGG | CCAGGCTTCG<br>GGTCCGAAGC  |                          |                           |                          |
| 10501 |                          | GCCTTTCTAC<br>CGGAAAGATG  |                          |                           |                          |
| 10551 |                          | GCATCTATCG<br>CGTAGATAGC  |                          |                           |                          |
| 10601 | GGCGCCCTCT<br>CCGCGGGAGA | TCCTCCCATG<br>AGGAGGGTAC  |                          |                           |                          |
| 10651 | AGCAGGGCTA<br>TCGTCCCGAT | GGTCGGCGAC<br>CCAGCCGCTG  | AACGCGCTCG<br>TTGCGCGAGC | GCTAATATGG<br>CGATTATACC  | CCTGCTGCAC<br>GGACGACGTG |
| 10701 |                          | GTAGACTGGA<br>CATCTGACCT  |                          |                           |                          |
| 10751 | CGCCCGTGTT               | GATGGTGTAA<br>CTACCACATT  | GTGCAGTTGG<br>CACGTCAACC | CCATAACGGA<br>GGTATTGCCT  | CCAGTTAACG<br>GGTCAATTGC |
| 10801 |                          | CCGGCTGCGA<br>GGCCGACGCT  |                          |                           |                          |
| 10851 |                          | AATACGTAGT<br>TTATGCATCA  |                          |                           |                          |
| 10901 |                          | GTGCGGCGGC<br>CACGCCGCCG  |                          |                           |                          |
| 10951 | GCCGGGGCTC<br>CGGCCCCGAG | CGGGGGCGAG<br>GCCCCCGCTC  | ATCTTCCAAC<br>TAGAAGGTTG | ATAAGGCGAT<br>TATTCCGCTA  | GATATCCGTA<br>CTATAGGCAT |
| 11001 | GATGTACCTG<br>CTACATGGAC | GACATCCAGG<br>CTGTAGGTCC  | TGATGCCGGC<br>ACTACGGCCG | GGCGGTGGTG<br>CCGCCACCAC  | GAGGCGCGCG<br>CTCCGCGCGC |
| 11051 | GAAAGTCGCG<br>CTTTCAGCGC | GACGCGGTTC<br>CTGCGCCAAG  | CAGATGTTGC<br>GTCTACAACG | GCAGCGGCAA<br>CGTCGCCGTT  | AAAGTGCTCC<br>TTTCACGAGG |
| 11101 | ATGGTCGGGA<br>TACCAGCCCT | CGCTCTGGCC<br>GCGAGACCGG  | GGTCAGGCGC<br>CCAGTCCGCG | GCGCAATCGT<br>CGCGTTAGCA  | TGACGCTCTA<br>ACTGCGAGAT |
| 11151 | GACCGTGCAA<br>CTGGCACGTT | AAGGAGAGCC<br>TTCCTCTCGG  | TGTAAGCGGG<br>ACATTCGCCC | CACTCTTCCG<br>GTGAGAAGGC  | TGGTCTGGTG<br>ACCAGACCAC |
| 11201 | CATAAATTCG<br>CTATTTAAGC | CAAGGGTATC<br>GTTCCCATAG  | ATGGCGGACG<br>TACCGCCTGC | ACCGGGGTTC<br>TGGCCCCAAG  | GAGCCCCGTA<br>CTCGGGGCAT |
| 11251 | TCCGGCCGTC<br>AGGCCGGCAG | CGCCGTGATC<br>CCGCGCACTAG | CATGCGGTTA<br>GTACGCCAAT | CCGCCCGCGT<br>CGCGGGGGGGA | GTCGAACCCA<br>CAGCTTGGGT |
| 11301 | GGTGTGCGAC<br>CCACACGCTG | GTCAGACAAC<br>CAGTCTGTTG  | CCCCTCACG                | TCCTTTTGGC<br>AGGAAAACCG  | TTCCTTCCAG<br>AAGGAAGGTC |

Figure 26L

|       | cececcecce<br>ecececeec  |      |                          |                          |
|-------|--------------------------|------|--------------------------|--------------------------|
| 11401 | AAGCGGTTAG<br>TTCGCCAATC |      | AGTGGCTCGC<br>TCACCGAGCG |                          |
| 11451 |                          |      | GGGACCCCCG<br>CCCTGGGGGC |                          |
| 11501 |                          |      | TTGCCTCCCC<br>AACGGAGGGG |                          |
| 11551 |                          |      | GACGAGCCCC<br>CTGCTCGGGG |                          |
| 11601 | · · · <del>-</del> · · - | <br> | CGCGGGGGGA               |                          |
| 11651 | GGCAAGAGCA<br>CCGTTCTCGT | <br> | GGGCACCCTC<br>CCCGTGGGAG |                          |
| 11701 |                          | <br> | GACGCGGCAG<br>CTGCGCCGTC |                          |
| 11751 |                          | <br> | CTACCTGGAC<br>GATGGACCTG |                          |
| 11801 |                          | <br> | CTCCTGAGCG<br>GAGGACTCGC | -                        |
| 11851 |                          | <br> |                          | GGCAGAACCT<br>CCGTCTTGGA |
| 11901 |                          |      |                          | GATCGAAAGT<br>CTAGCTTTCA |
| 11951 |                          |      | TGAATCGCGA<br>ACTTAGCGCT | GCGGTTGCTG<br>CGCCAACGAC |
| 12001 |                          | <br> | ACCGGGATTA<br>TGGCCCTAAT | GTCCCGCGCG<br>CAGGGCGCGC |
| 12051 |                          |      |                          | CAGACGGTGA<br>GTCTGCCACT |
| 12101 |                          |      |                          | GCGTACGCTT<br>CGCATGCGAA |
| 12151 |                          |      |                          | GGGACTTTGT<br>CCCTGAAACA |
| 12201 |                          |      |                          | GCGCAGCTGT<br>CGCGTCGACA |
| 12251 |                          |      |                          | GGATGCGCTG<br>CCTACGCGAC |

7 igure 26 M

| 12301 | CTAAACATAG<br>GATTTGTATC | T GCCCGA<br>ATCTCGGGCT   | GGGCGCGACC               | CTGCTCGATT<br>GACGAGCTAA | TGATAA TT<br>ACTATTTGTA      |
|-------|--------------------------|--------------------------|--------------------------|--------------------------|------------------------------|
| 12351 | CCTGCAGAGC<br>GGACGTCTCG | ATAGTGGTGC<br>TATCACCACG | AGGAGCGCAG<br>TCCTCGCGTC | CTTGAGCCTG<br>GAACTCGGAC | GCTGACAAGG<br>CGACTGTTCC     |
| 12401 | TGGCCGCCAT<br>ACCGGCGGTA | CAACTATTCC<br>GTTGATAAGG | ATGCTTAGCC<br>TACGAATCGG | TGGGCAAGTT<br>ACCCGTTCAA | TTACGCCCGC<br>AATGCGGGCG     |
| 12451 | AAGATATACC<br>TTCTATATGG | ATACCCCTTA<br>TATGGGGAAT | CGTTCCCATA<br>GCAAGGGTAT | GACAAGGAGG<br>CTGTTCCTCC | TAAAGATCGA<br>ATTTCTAGCT     |
| 12501 | GGGGTTCTAC<br>CCCCAAGATG | ATGCGCATGG<br>TACGCGTACC | CGCTGAAGGT<br>GCGACTTCCA | GCTTACCTTG<br>CGAATGGAAC | AGCGACGACC<br>TCGCTGCTGG     |
| 12551 | ACCCGCAAAT               | AGCGTTGCTC               | GCGTAGGTGT               |                          | GCACTCGGCC                   |
| 12601 |                          | AGTCGCTGGC               | GCTCGACTAC               | GTGTCGGACG               | TTTCCCGGGA                   |
| 12651 | CCGACCGTGC               | CCGTCGCCGC               | TATCTCTCCG               | GCTCAGGATG               | TTTGACGCGG<br>AAACTGCGCC     |
| 12701 | CGCGACTGGA               | CGCGACCCGG               | GGTTCGGCTG               | CGCGGGACCT               | GGCAGCTGGG<br>CCGTCGACCC     |
| 12751 | CGGCCTGGAC               | CCGACCGCCA               | CCGTGGGCGC               | GCGCGACCGT               | ACGTCGGCGG<br>TGCAGCCGCC     |
| 12801 | CGTGGAGGAA<br>GCACCTCCTT | TATGACGAGG<br>ATACTGCTCC | ACGATGAGTA<br>TGCTACTCAT | CGAGCCAGAG               | GACGGCGAGT<br>CTGCCGCTCA     |
| 12851 | TGATTCGCCA               | CTACAAAGAC               | TAGTCTACTA               | CGTTCTGCGT               | ACGGACCCGG                   |
| 12901 | GCCACGCCCG               | CCGCGACGTC               | TCGGTCGGCA               | GGCCGGAATT               | CTCCACGGAC<br>GAGGTGCCTG     |
| 12951 |                          | TCCAGTACCT               | GGCGTAGTAC               | AGCGACTGAC               | GCGCGTTAGG                   |
| 13001 | ACTGCGCAAG               | GCCGTCGTCG               | GCGTCCGGTT               | GGCCGAGAGG               | GCAATTCTGG<br>CGTTAAGACC     |
|       | TTCGCCACCA               | GGGCCGCGCG               | CGTTTGGGGT               | GCGTGCTCTT               | GCTGCTGGCG<br>CCACGACCGC     |
| 13101 | ATCGTAAACG<br>TAGCATTTGC | CGCTGGCCGA               | AAACAGGGCC               | TAGGCCGGGC               | : ACGAGGCCGG<br>: TGCTCCGGCC |
| 13151 | CCTGGTCTAC<br>GGACCAGATG | GACGCGCTGC<br>CTGCGCGACG | TTCAGCGCGT               | GGCTCGTTAC               | AACAGCGGCA<br>TTGTCGCCGT     |
| 13201 | ACGTGCAGAC<br>TGCACGTCTG | CAACCTGGAC<br>GTTGGACCTC | CGGCTGGTGG<br>CCCGACCACC | GGGATGTGCG<br>CCCTACACGC | CGAGGCCGTG<br>CCTCCGGCAC     |

Figure 26 N.

| 13251  | GCGCAGCGTG<br>CGCGTCGCAC | A CGCGCGCA<br>TCGCGCGCGT | GCAGCAGGGC<br>CGTCGTCCCG | AACCTGGGCT<br>TTGGACCCGA | CCATGG C<br>GGTACCAACG   |
|--------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| 13301  |                          |                          |                          | CAACGTGCCG<br>GTTGCACGGC |                          |
| 13351  |                          |                          |                          | GGCTAATGGT<br>CCGATTACCA |                          |
| 13401  |                          |                          |                          | GACTATTTT<br>CTGATAAAAA  |                          |
| 13451  |                          |                          |                          | CCAGGCTTTC<br>GGTCCGAAAG |                          |
| 13501  |                          |                          |                          | GCGACCGCGC<br>CGCTGGCGCG |                          |
| 13551  |                          |                          |                          | CTGCTGCTAA<br>GACGACGATT |                          |
| 13601  |                          |                          |                          | ATACCTAGGT<br>TATGGATCCA |                          |
| 13651  |                          |                          |                          | ATGTGGACGA<br>TACACCTGCT |                          |
| 13701  |                          |                          |                          | GGGCAGGAGG<br>CCCGTCCTCC |                          |
| 13751  |                          |                          |                          | CAACCGGCGG<br>GTTGGCCGCC |                          |
| 13801  |                          |                          |                          | AGCGCATTTT<br>TCGCGTAAAA |                          |
| 13851  |                          |                          |                          | GACGGGGTAA<br>CTGCCCCATT | CGCCCAGCGT<br>GCGGGTCGCA |
| 13901  |                          |                          |                          | ACCGGGCATG<br>TGGCCCGTAC | TATGCCTCAA<br>ATACGGAGTT |
| 13,951 |                          |                          | •                        | ACTTGCATCG<br>TGAACGTAGC | CCCCCCCCC                |
| 14001  |                          |                          |                          |                          | ACTGGCTACC<br>TGACCGATGG |
| 14051  |                          |                          |                          |                          | GGTAACGATG<br>CCATTGCTAC |
| 14101  |                          |                          |                          |                          | GCAACCGCAG<br>CGTTGGCGTC |
| 14151  |                          |                          |                          |                          | CGCTGCGAAA<br>GCGACGCTTT |

7, gure 260

| 14201 | GGAAAGCTTC               | COGAGGCCAA               | GCAGCTTGTC<br>CGTCGAACAG | CGATCTAGGC*<br>GCTAGATCCG | GCTGCGCCCC<br>CGACGC     |
|-------|--------------------------|--------------------------|--------------------------|---------------------------|--------------------------|
|       | CGCGGTCAGA               |                          |                          |                           |                          |
| 14251 | GCGCCAGTCT               | ACGATCATCG               | GGTAAAGGTT               | CGAACTATCC                | CAGAGAATGG               |
| 14301 | AGCACTCGCA               | CCACCGGCCC               | GCGCCTGCTG               | GGCGAGGAGG                | AGTACCTAAA               |
|       |                          |                          | CGCGGACGAC               |                           | ·                        |
| 14351 | CAACTCGCTG<br>GTTGAGCGAC | CTGCAGCCGC<br>GACGTCGGCG | AGCGCGAAAA<br>TCGCGCTTTT | AAACCTGCCT<br>TTTGGACGGA  | CCGGCATTTC<br>GGCCGTAAAG |
| 14401 | CCAACAACGG               | GATAGAGAGC               | CTAGTGGACA               | AGATGAGTAG                | ATGGAAGACG               |
| 14401 | GGTTGTTGCC               | CTATCTCTCG               | GATCACCTGT               | TCTACTCATC                | TACCTTCTGC               |
| 14451 | TACGCGCAGG               | AGCACAGGGA               | CGTGCCAGGC               | CCGCGCCCGC                | CCACCCGTCG               |
|       | ATGCGCGTCC               | TCGTGTCCCT               | GCACGGTCCG               | GGCGCGGCG                 | GGTGGGCAGC               |
| 14501 | TCAAAGGCAC               | GACCGTCAGC               | GGGGTCTGGT               | GTGGGAGGAC                | GATGACTCGG               |
|       |                          |                          |                          |                           | CTACTGAGCC               |
| 14551 | CAGACGACAG               | CAGCGTCCTG               | GATTTGGGAG               | GGAGTGGCAA                | CCCGTTTGCG               |
|       |                          |                          |                          |                           | GGGCAAACGC               |
| 14601 | CACCTTCGCC               | CCAGGCTGGG               | GAGAATGTTT               | TAAAAAAAAA                | AAAAGCATGA               |
|       |                          |                          |                          |                           | TTTTCGTACT               |
| 14651 | TGCAAAATAA               | AAAACTCACC               | AAGGCCATGG               | CACCGAGCGT                | TGGTTTTCTT               |
|       |                          |                          |                          |                           | ACCAAAAGAA               |
| 14701 | GTATTCCCCT               | TAGTATGCGG               | CGCGCGGCGA               | . TGTATGAGGA              | AGGTCCTCCT               |
|       |                          |                          |                          |                           | TCCAGGAGGA               |
| 14751 | CCCTCCTACG               | AGAGTGTGGI               | GAGCGCGGCG               | CCAGTGGCGG                | CGGCGCTGGG               |
|       |                          |                          |                          |                           | GCCGCGACCC               |
| 14801 | TTCTCCCTTC               | GATGCTCCCC               | TGGACCCGCC               | GTTTGTGCCT                | CCGCGGTACC               |
|       |                          |                          |                          |                           | GGCGCCATGG               |
| 14851 | TGCGGCCTAC               | CGGGGGGAGA               | AACAGCATCC               | GTTACTCTGA                | GTTGGCACCC               |
|       |                          |                          |                          |                           | CAACCGTGGG               |
| 14901 | CTATTCGACA               | CCACCCGTGT               | GTACCTGGTG               | GACAACAAGT                | CAACGGATGT               |
|       |                          |                          |                          |                           | A GTTGCCTACA             |
| 14951 | GGCATCCCTG               | AACTACCAG                | ACGACCACA                | CAACTTTCT                 | ACCACGGTCA               |
|       |                          |                          |                          |                           | TGGTGCCAGT               |
| 15001 | TTCAAAACAA               | TGACTACAG                | CCGGGGGAG(               | CAAGCACAC                 | GACCATCAAT               |
|       |                          |                          |                          |                           | CTGGTAGTTA               |
| 15051 | CTTGACGACC               | GGTCGCACT                | GGGCGGCGA                | CTGAAAACC                 | TCCTGCATAC               |
|       |                          |                          |                          |                           | r AGGACGTATG             |
| 15101 | CAACATGCC                | AATGTGAAC                | AGTTCATGT                | TACCAATAA                 | TTTAAGGCGC               |
|       | GTTGTACGGT               | TTACACTTG                | C TCAAGTACA              | A ATGGTTATT               | AAATTCCGCG               |

Figure 26 P

| 15151  |            | CCCTTG                   |            |            |                          |
|--------|------------|--------------------------|------------|------------|--------------------------|
|        | CCCACTACCA | CAGCGCGAAC               | GGATGATTCC | TGTTAGTCCA | CCTCGACTTT               |
| 15201  |            | TGGAGTTCAC<br>ACCTCAAGTG |            |            |                          |
| 15251  |            | CTTATGAACA<br>GAATACTTGT |            |            |                          |
| 15301  |            | CGGGGTTCTG<br>GCCCCAAGAC |            |            |                          |
| 15351  |            | GACTGGGGTT<br>CTGACCCCAA |            |            |                          |
| 15401  |            | AACGAAGCCT<br>TTGCTTCGGA |            |            |                          |
| 15451  |            | CTTCACCCAC<br>GAAGTGGGTG |            |            |                          |
| 15501  | AAGCGGCAAC | CCTTCCAGGA<br>GGAAGGTCCT | GGGCTTTAGG | ATCACCTACG | ATGATCTGGA               |
|        |            |                          |            |            |                          |
| 15551  |            | ATTCCCGCAC<br>TAAGGGCGTG |            | •          |                          |
| 15601  |            | CACCGAACAG<br>GTGGCTTGTC |            |            |                          |
| 15651  |            | GCGCGGAAGA<br>CGCGCCTTCT |            |            |                          |
| 15701  |            | GACATGAACG<br>CTGTACTTGC |            |            |                          |
| 15751  |            | GGAGAAGCGC<br>CCTCTTCGCG |            |            |                          |
| 15801  |            | CGCAACCCGA<br>GCGTTGGGCT |            |            |                          |
| 15851  |            | ACAGAGGACA<br>TGTCTCCTGT |            |            | CTAATAAGCA<br>GATTATTCGT |
|        |            |                          |            |            | ATACAACTAC<br>TATGTTGATG |
| 15951  |            |                          |            |            | GCACTCCTGA<br>CGTGAGGACT |
| 160,01 |            | •                        |            |            | GACATGATGC<br>CTGTACTACG |
| 16051  |            |                          |            |            | CTTTCCGGTG<br>GAAAGGCCAC |

Figure 26 Q

| 16101 | GTGGGCGCCG<br>CACCCGCGGC | A TGTTGCC<br>T ACAACGG   | CGTGCACTCC<br>GCACGTGAGG                | AAGAGCTTCT<br>TTCTCGAAGA | ACAACGA CA<br>TGTTGC GT   |
|-------|--------------------------|--------------------------|---|--------------------------|---------------------------|
| 16151 | GGCCGTCTAC<br>CCGGCAGATG | TCCCAACTCA<br>AGGGTTGAGT | TCCGCCAGTT<br>AGGCGGTCAA                | TACCTCTCTG<br>ATGGAGAGAC | ACCCACGTGT<br>TGGGTGCACA  |
| 16201 | TCAATCGCTT<br>AGTTAGCGAA | TCCCGAGAAC<br>AGGGCTCTTG | CAGATTTTGG<br>GTCTAAAACC                | ececececee<br>Cececccecc | AGCCCCCACC<br>TCGGGGGTGG  |
| 16251 | ATCACCACCG<br>TAGTGGTGGC |                          |   | CTCACAGATC<br>GAGTGTCTAG |                           |
| 16301 | ACCGCTGCGC<br>TGGCGACGCG | AACAGCATCG<br>TTGTCGTAGC | GAGGAGTCCA<br>CTCCTCAGGT                | GCGAGTGACC<br>CGCTCACTGG | ATTACTGACG<br>TAATGACTGC  |
| 16351 | CCAGACGCCG<br>GGTCTGCGGC | CACCTGCCCC<br>GTGGACGGGG | TACGTTTACA<br>ATGCAAATGT                | AGGCCCTGGG<br>TCCGGGACCC | CATAGTCTCG<br>GTATCAGAGC  |
| 16401 | CCGCGCGTCC               | TATCGAGCCG<br>ATAGCTCGGC | CACTTTTGA<br>GTGAAAAACT                 | GCAAGCATGT<br>CGTTCGTACA | CCATCCTTAT<br>GGTAGGAATA  |
| 16451 | ATCGCCCAGC<br>TAGCGGGTCG | AATAACACAG<br>TTATTGTGTC | GCTGGGGCCT<br>CGACCCCGGA                | GCGCTTCCCA<br>CGCGAAGGGT | AGCAAGATGT<br>TCGTTCTACA  |
| 16501 | TTGGCGGGGC<br>AACCGCCCG  | CAAGAAGCGC<br>GTTCTTCGCG | TCCGACCAAC<br>AGGCTGGTTG                | ACCCAGTGCG<br>TGGGTCACGC | 222222222<br>222222222    |
| 16551 | CACTACCGCG<br>GTGATGGCGC | CGCCCTGGGG<br>GCGGGACCCC | CGCGCACAAA<br>GCGCGTGTTT                | CGCGGCCGCA               | CTGGGCGCAC<br>GACCCGCGTG  |
| 16601 |                          |                          |   |                          | CGCAACTACA<br>GCGTTGATGT  |
| 16651 | CGCCCACGCC               | GCCACCAGTG<br>CGGTGGTCAC | TCCACAGTGG<br>AGGTGTCACC                | ACGCGGCCAT<br>TGCGCCGGTA | TCAGACCGTG<br>AGTCTGGCAC  |
| 16701 | GTGCGCGGAG<br>CACGCGCCTC | CCCGGCGCTA<br>GGGCCGCGAT | TGCTAAAATG<br>ACGATTTTAC                | AAGAGACGGC<br>TTCTCTGCCG | GGAGGCGCGT<br>CCTCCGCGCA  |
| 16751 | AGCACGTCGC<br>TCGTGCAGCG | CACCGCCGCC<br>GTGGCGGCGG | GACCCGGCAC<br>CTGGGCCGTG                | TGCCGCCCAA<br>ACGGCGGGTT | CGCGCGCGCG                |
| 16801 | CGGCCCTGCT<br>GCCGGGACGA | TAACCGCGCA<br>ATTGGCGCGT | CGTCGCACCG                              | GCCGACGGGC               | GGCCATGCGG                |
| 16851 | GCCGCTCGAA<br>CGGCGAGCTT | GGCTGGCCGC               | GGGTATTGTC                              | ACTGTGCCCC<br>TGACACGGGG | CCAGGTCCAG                |
| 16901 | GCGACGAGCG<br>CGCTGCTCGC | CCCCCCCAC                | CAGCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC | CATTAGTGCT<br>GTAATCACGA | ATGACTCAGG<br>TACTGAGTCC  |
| 16951 | GTCGCAGGGG<br>CAGCGTCCCC | CAACGTGTAT<br>GTTGCACATA | TGGGTGCGCG                              | ACTCGGTTAG               | CGCCCTGCGC:<br>CCCGGACGCG |
| 17001 | GTGCCCGTGC<br>CACGGGCACG | GCACCGCCC<br>CGTGGGCGGG  | CCCGCGCAAC                              | TAGATTGCAA<br>ATCTAACGTT | GAAAAAACTA<br>CTTTTTTGAT  |



| 17051  |                          |                          |                          | cceccéccec<br>eccececene |                          |
|--------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| 17101  |                          |                          |                          | TGCTCCAGGT<br>ACGAGGTCCA |                          |
| 17151  |                          |                          |                          | CAGGATTACA<br>GTCCTAATGT |                          |
| 17201  |                          | _                        |                          | TGATGATGAT<br>ACTACTACTA |                          |
| 17251  |                          |                          |                          | CCAGGCGACG<br>GGTCCGCTGC |                          |
| 17301  |                          |                          |                          | CCCGGCACCA<br>GGGCCGTGGT |                          |
| 17351  |                          |                          |                          | CAAGCGCGTG<br>GTTCGCGCAC |                          |
| 17401  |                          |                          |                          | CCAACGAGCG<br>GGTTGCTCGC |                          |
| 17451  |                          |                          |                          | CTGGCGTTGC<br>GACCGCAACG |                          |
| 17501  |                          |                          |                          | AACACTGCAG<br>TTGTGACGTC |                          |
| 17551  |                          |                          |                          | GCCTAAAGCG<br>CGGATTTCGC |                          |
| .17601 |                          |                          |                          | CCCAAGCGCC<br>GGGTTCGCGG |                          |
| 17651  |                          |                          |                          | TGGGCTGGAG<br>ACCCGACCTC |                          |
| 17701  |                          |                          |                          | GACTGGGCGT<br>CTGACCCGCA |                          |
| 17751  |                          |                          |                          |                          | CCGCCACAGA<br>GGCGGTGTCT |
| 17801  |                          |                          |                          |                          | GCGGATGCCG<br>CGCCTACGGC |
| 17851  | CGGTGCAGGC<br>GCCACGTCCG | GCTCGCTGCG<br>CCAGCGACGC | GCCGCGTCCA<br>CGGCGCAGGT | AGACCTCTAC<br>TCTGGAGATG | GGAGGTGCAA<br>CCTCCACGTT |
| 17901  | ACGGACCCGT<br>TGCCTGGGCA | GGATGTTTCG<br>CCTACAAAGC | CGTTTCAGCC<br>GCAAAGTCGG | CCCCCCCCC                | CGCGCCGTTC<br>GCGCGGCAAG |
| 17951  |                          |                          |                          |                          | GCCCTACATC<br>CGGGATGTAG |

Figure 265

| 18001  | CTTCCATTGC<br>GAAGGTAACG | GCCTACCCCC<br>CTATGGGGG  | GGCTATCGTG<br>CCGATAGCAC     | GCTACACCTAL<br>CGATGTGGAT | CCCCCC T                 |
|--------|--------------------------|--------------------------|------------------------------|---------------------------|--------------------------|
| 18051  | AGACGAGCAA<br>TCTGCTCGTT | CTACCCGACG<br>GATGGGCTGC | CCGAACCACC<br>GGCTTGGTGG     | ACTGGAACCC<br>TGACCTTGGG  | CCCCCCCCCC               |
| 18101  | TCGCCGTCGC               | CAGCCCGTGC               | TGGCCCCGAT                   | TTCCGTGCGC                | AGGGTGGCTC               |
|        | AGCGGCAGCG               | GTCGGGCACG               | ACCGGGGCTA                   | AAGGCACGCG                | TCCCACCGAG               |
| 18151  | GCGAAGGAGG               | CAGGACCCTG               | GTGCTGCCAA                   | CAGCGCGCTA                | CCACCCCAGC               |
|        | CGCTTCCTCC               | GTCCTGGGAC               | CACGACGGTT                   | GTCGCGCGAT                | GOTGGGGTCG               |
| 18201  | ATCGTTTAAA               | AGCCGGTCTT               | TGTGGTTCTT                   | GCAGATATGG                | CCCTCACCTG               |
|        | TAGCAAATTT               | TCGGCCAGAA               | ACACCAAGAA                   | CGTCTATACC                | GGGAGTGGAC               |
| 18251  | CCGCCTCCGT               | TTCCCGGTGC               | CGGGATTCCG                   | AGGAAGAATG                | CACCGTAGGA               |
|        | GGCGGAGGCA               | AAGGGCCACG               | GCCCTAAGGC                   | TCCTTCTTAC                | GTGGCATCCT               |
| 18301  | GGGGCATGGC<br>CCCCGTACCG | CGGCCACGGC               | CTGACGGGCG<br>GACTGCCCGC     | GCATGCGTCG<br>CGTACGCAGC  | TGCGCACCAC<br>ACGCGTGGTG |
| 18351  | CCCCCCCCC                | GCGCGCAGCGT              | CCGTCGCATG<br>GGCAGCGTAC     | CGCGGCGGTA<br>GCGCCGCCAT  | TCCTGCCCCT<br>AGGACGGGGA |
| 18401  | CCTTATTCCA<br>GGAATAAGGT | CTGATCGCCG<br>GACTAGCGGC | CGGCGATTGG<br>GCCGCTAACC     | CGCCGTGCCC                | GGAATTGCAT<br>CCTTAACGTA |
| 18451  | CCGTGGCCTT               | GCAGGCGCAG               | AGACACTGAT                   | TAAAAACAAG                | TTGCATGTGG               |
|        | GGCACCGGAA               | CGTCCGCGTC               | TCTGTGACTA                   | ATTTTTGTTC                | AACGTACACC               |
| 18501  | AAAAATCAAA               | ATAAAAAGTC               | TGGACTCTCA                   | CGCTCGCTTG                | GTCCTGTAAC               |
|        | TTTTTAGTTT               | TATTTTTCAG               | ACCTGAGAGT                   | GCGAGCGAAC                | CAGGACATTG               |
| 18551  | TATTTTGTAG<br>ATAAAACATC | AATGGAAGAC<br>TTACCTTCTG | ATCAACTTTG<br>TAGTTGAAAC     | CGTCTCTGGC                | CCCGCGACAC<br>GGGCGCTGTG |
| 18,601 | GGCTCGCGCC               | CGTTCATGGG               | AAACTGGCAA                   | GATATCGGCA                | CCAGCAATAT               |
|        | CCGAGCGCGG               | GCAAGTACCC               | TTTGACCGTT                   | CTATAGCCGT                | GGTCGTTATA               |
| 18651  | GAGCGGTGGC               | GCCTTCAGCT               | GGGGCTCGCT                   | GTGGAGCGGC                | TTAAAAATTA               |
|        | CTCGCCACCG               | GCGAAGTCGA               | CCCCGAGCGA                   | CACCTCGCCG                | AATTTTTAAT               |
| 18701  | TCGGTTCCAC               | CGTTAAGAAC               | TATGGCAGGA                   | AGGCCTGGAA                | CAGCAGCACA               |
|        | AGCCAAGGTG               | GCAATTCTTC               | ATACCGTCGT                   | TCCGGACCTT                | GTCGTCGTGT               |
| 18751  | GGCCAGATGC<br>CCGGTCTACG | TGAGGGATAF<br>ACTCCCTATT | CAACTITCTC                   | CAAAATTTCC<br>CTTTTAAAGG  | AACAAAAGGT<br>TTGTTTTCCA |
| 18801  | GGTAGATGGC               | CIGGCCTCTC               | GCATTAGCGC                   | GGTGGTGGAC                | CTGGCCAACC               |
|        | CCATCTACCC               | GACCGGAGAC               | CGTAATCGCC                   | CCACCACCTG                | GACCGGTTGG               |
| 18851  | AGGCAGTGCA               | AAATAAGATI               | AACAGTAAGO                   | TTGATCCCCG                | CCCTCCCGTA               |
|        | TCCGTCACGT               | TTTATTCTA                | A TTGTCATTCO                 | AACTAGGGGC                | GGGAGGGCAT               |
| 18901  | GAGGAGCCTC<br>CTCCTCGGAG | CACCGGCCGT               | r GGAGACAGTO<br>A CCTCTGTCAO | TCTCCAGAGG<br>AGAGGTCTCC  | GGCGTGGCGA<br>CCGCACCGCT |

Figure 26T

| 18951 | AAAGCGTCCG<br>TTTCGCAGGC |      | TCTGGTGACG<br>AGACCACTGC  |                          |
|-------|--------------------------|------|---------------------------|--------------------------|
| 19001 |                          |      | AAGGCCTGCC<br>TTCCGGACGG  |                          |
| 19051 |                          |      | GGCCAGCACA<br>CCGGTCGTGT  |                          |
| 19101 |                          |      | GCAGAAACCT<br>CGTCTTTGGA  |                          |
| 19151 |                          | <br> | GCCGCGCGTC<br>CGGCGCGCAG  |                          |
| 19201 |                          |      | GTAGCCAGTG<br>CATCGGTCAC  |                          |
| 19251 |                          |      | GGTGCAATCC<br>CCACGTTAGG  |                          |
| 19301 | GACGATGCTT<br>CTGCTACGAA |      | TGTGTGTCAT<br>ACACACAGTA  |                          |
| 19351 |                          | <br> | CGCGCGCCCG                |                          |
| 19401 |                          |      | CTTACATGCA<br>GAATGTACGT  |                          |
| 19451 |                          |      | CTGGTGCAGT.<br>GACCACGTCA |                          |
| 19501 |                          |      | GTTTAGAAAC<br>CAAATCTTTG  |                          |
| 19551 |                          | <br> | CCCAGCGTTT<br>GGGTCGCAAA  |                          |
| 19601 |                          |      | TACTCGTACA<br>ATGAGCATGT  |                          |
| 19651 |                          | <br> | GGACATGGCT<br>CCTGTACCGA  |                          |
| 19701 | TTGACATCCG<br>AACTGTAGGC |      |                           | GCCCTACTCT<br>CGGGATGAGA |
| 19751 | GGCACTGCCT<br>CCGTGACGGA |      | GGTGCCCCAA<br>CCACGGGGTT  |                          |
| 19801 | ATGGGATGAA<br>TACCCTACTT |      | AAACCTAGAA<br>TTTGGATCTT  |                          |
| 19851 |                          |      |                           | AAAAACTCAC<br>TTTTTGAGTG |

Figure 26 U

| 19901 | GTATTTGGGC<br>CATAAACCCG | A GCCTTA<br>TCCGCGGAAT   | TTCTGGTATA<br>AAGACCATAT | AATATTACAA<br>TTATAATGTT | AGGAGG T<br>TCCTCCCATA   |
|-------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| 19951 |                          | GTCGAAGGTC<br>CAGCTTCCAG |                          |                          |                          |
| 20001 |                          | TCAAATAGGA<br>AGTTTATCCT |                          |                          |                          |
| 20051 |                          | GGAGAGTCCT<br>CCTCTCAGGA |                          |                          |                          |
| 20101 |                          | GCAAAACCCA<br>CGTTTTGGGT |                          |                          |                          |
| 20151 | TARAGCAACA<br>ATTTCGTTGT | AAATGGAAAG<br>TTTACCTTTC | CTAGAAAGTC<br>GATCTTTCAG | AAGTGGAAAT<br>TTCACCTTTA | GCAATTTTTC<br>CGTTAAAAAG |
| 20201 |                          | AGGCAGCCGC<br>TCCGTCGGCG |                          |                          |                          |
| 20251 |                          | AGTGAAGATG<br>TCACTTCTAC |                          |                          |                          |
| 20301 |                          | CACTATTAAG<br>GTGATAATTC |                          |                          |                          |
| 20351 |                          | CCAACAGGCC<br>GGTTGTCCGG |                          |                          |                          |
| 20401 |                          | TATTACAACA<br>ATAATGTTGT |                          |                          |                          |
| 20451 |                          | GTTGAATGCT<br>CAACTTACGA |                          |                          |                          |
| 20501 |                          |                          |                          |                          | CCAGGTACTT<br>GGTCCATGAA |
| 20551 |                          |                          |                          |                          | GTTAGAATTA<br>CAATCTTAAT |
| 20601 |                          |                          |                          |                          | CTTTCCACTG<br>GAAAGGTGAC |
| 20651 | GGAGGTGTGA<br>CCTCCACACT | TTAATACAGA<br>AATTATGTCT | GACTCTTACC<br>CTGAGAATGG | AAGGTAAAAC<br>TTCCATTTTG | CTAAAACAGG<br>GATTTTGTCC |
| 20701 | TCAGGAAAAT<br>AGTCCTTTTA | GGATGGGAAA<br>CCTACCCTTT | AAGATGCTAC<br>TTCTACGATG | AGAATTTTCA<br>TCTTAAAAGT | GATAAAAATG<br>CTATTTTTAC |
| 20751 | AAATAAGAGT<br>TTTATTCTCA | TGGAAATAAT<br>ACCTTTATTA | TTTGCCATGG<br>AAACGGTACC | AAATCAATCT<br>TTTAGTTAGA | AAATGCCAAC<br>TTTACGGTTG |
| 20801 | CTGTGGAGAA<br>GACACCTCTT | ATTTCCTGTA<br>TAAAGGACAT | CTCCAACATA<br>GAGGTTGTAT | GCGCTGTATT<br>CGCGACATAA | TGCCCGACAA<br>ACGGGCTGTT |

Figure 26 V

| 20851 | GCTAAAGTAC<br>CGATTTCATG | ACCTTCCA<br>TGAAGGT      | ACGTAAAAAT<br>TGCATTTTTA | TTCTGATÄÄČ<br>AAGACTATTG | TCCAÄACTOTT<br>GCTTTGCA  |
|-------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| 20901 | ACGACTACAT<br>TGCTGATGTA | GAACAAGCGA<br>CTTGTTCGCT | GTGGTGGCTC<br>CACCACCGAG | CCGGGCTAGT<br>GGCCCGATCA | GGACTGCTAC<br>CCTGACGATG |
| 20951 |                          |                          |                          | TATATGGACA<br>ATATACCTGT |                          |
| 21001 | ATTTAACCAC<br>TAAATTGGTG |                          |                          | CTACCGCTCA<br>GATGGCGAGT |                          |
| 21051 |                          |                          |                          | AGGTGCCTCA<br>TCCACGGAGT |                          |
| 21101 | ••••                     |                          |                          | TCATACACCT<br>AGTATGTGGA |                          |
| 21151 |                          |                          |                          | GAGCTCCCTA<br>CTCGAGGGAT |                          |
| 21201 |                          |                          |                          | ATAGCATTIG<br>TATCGTAAAC |                          |
| 21251 |                          |                          |                          | TCCACGCTTG<br>AGGTGCGAAC |                          |
| 21301 |                          |                          |                          | CGACTATCTC<br>GCTGATAGAG |                          |
| 21351 |                          |                          |                          | CCAACGTGCC<br>GGTTGCACGG |                          |
| 21401 |                          |                          |                          | TGGGCCTTCA<br>ACCCGGAAGT |                          |
| 21451 |                          |                          |                          | CTACGACCCT<br>GATGCTGGGA |                          |
| 21501 |                          |                          |                          | CCTTTTACCT<br>GGAAAATGGA |                          |
| 21551 |                          |                          |                          | TCTGTCAGCT<br>AGACAGTCGA |                          |
| 21601 | TGACCGCCTG<br>ACTGGCGGAC | CTTACCCCCA<br>GAATGGGGGT | ACGAGTTTGA<br>TGCTCAAACT | AATTAAGCGC<br>TTAATTCGCG | TCAGTTGACG<br>AGTCAACTGC |
| 21651 | GGGAGGGTTA<br>CCCTCCCAAT | CAACGTTGCC<br>GTTGCAACGG | CAGTGTAACA<br>GTCACATTGT | TGACCAAAGA<br>ACTGGTTTCT | CTGGTTCCTG<br>GACCAAGGAC |
| 21701 | GTACAAATGC<br>CATGTTTACG | TAGCTAACTA<br>ATCGATTGAT | TAACATTGGC<br>ATTGTAACCG | TACCAGGGCT<br>ATGGTCCCGA | TCTATATCCC<br>AGATATAGGG |
| 21751 | AGAGAGCTAC<br>TCTCTCGATG | AAGGACCGCA<br>TTCCTGGCGT | TGTACTCCTT<br>ACATGAGGAA | CTTTAGAAAC<br>GAAATCTTTG | TTCCAGCCCA<br>AAGGTCGGGT |

Figure 26 W

| 21801 | TGAGCCGTCA<br>ACTCGGCAGT | CONCETTS CALCUTA         | GATACŤAAAT<br>CTATGATTTA | ACAAGGACTAS<br>TGTTCCTGAT | GGTTGT C                 |
|-------|--------------------------|--------------------------|--------------------------|---------------------------|--------------------------|
| 21851 |                          |                          |                          | TTTGTTGGCT<br>AAACAACCGA  |                          |
| 21901 |                          |                          |                          | TAACTTCCCC<br>ATTGAAGGGG  |                          |
| 21951 | TAGGCAAGAC<br>ATCCGTTCTG | CGCAGTTGAC<br>GCGTCAACTG | AGCATTACCC<br>TCGTAATGGG | AGAAAAAGTT<br>TCTTTTTCAA  | TCTTTGCGAT<br>AGAAACGCTA |
| 22001 |                          |                          |                          | AACTTTATGT<br>TTGAAATACA  |                          |
| 22051 | ACTCACAGAC<br>TGAGTGTCTG | CTGGGCCAAA<br>GACCCGGTTT | ACCTTCTCTA<br>TGGAAGAGAT | CGCCAACTCC<br>GCGGTTGAGG  | GCCCACGCGC               |
| 22101 | TAGACATGAC<br>ATCTGTACTG | TTTTGAGGTG<br>AAAACTCCAC | GATCCCATGG<br>CTAGGGTACC | ACGAGCCCAC<br>TGCTCGGGTG  | CCTTCTTTAT<br>GGAAGAAATA |
| 22151 | GTTTTGTTTG<br>CAAAACAAAC | AAGTCTTTGA<br>TTCAGAAACT | CGTGGTCCGT<br>GCACCAGGCA | GTGCACCAGC<br>CACGTGGTCG  | CGCACCGCGG<br>GCGTGGCGCC |
| 22201 | GCAGTAGCTT               | TGGCACATGG               | ACGCGTGCGG               | CTTCTCGGCC<br>GAAGAGCCGG  | CCGTTGCGGT               |
| 22251 |                          |                          |                          | AGCTGCCGCC<br>TCGACGGCGG  |                          |
| 22301 | CACTCGTCCT               | TGACTTTCGG               | TAACAGTTTC               | ATCTTGGTTG<br>TAGAACCAAC  | ACCCGGTATA               |
| 22351 | AAAAACCCGT               | GGATACTGTT               | CGCGAAAGGT               | GGCTTTGTTT                | GAGGTGTGTT               |
| 22401 | GCTCGCCTGC<br>CGAGCGGACG | GCCATAGTCA<br>CGGTATCAGT | ATACGGCCGG<br>TATGCCGGCC | TCGCGAGACT<br>AGCGCTCTGA  | GGGGGCGTAC<br>CCCCGCATG  |
| 22451 |                          |                          |                          | CAAAAACATG<br>GTTTTTGTAC  |                          |
| 22501 | CTCGGGAAAC               | CGAAAAGACT               | GGTCGCTGAG               | AAGCAGGTTT<br>TTCGTCCAAA  | TGGTCAAACT               |
|       | CATGCTCAGT               | GAGGACGCGG               | CATCGCGGTA               | ACGAAGAAGG                | CCCGACCGCT               |
|       | CATATTGCGA               | CCTTTTCAGG               | TGGGTTTCGC               | ATGTCCCCGG                | CAACTCGGCC<br>GTTGAGCCGG |
| 22651 | GCCTGTGGAC<br>CGGACACCTG | TATTCTGCTG<br>ATAAGACGAC | CATGTTTCTC               | CACGCCTTTG<br>GTGCGGAAAC  | CCAACTGGCC<br>GGTTGACCGG |
| 22701 | CCAAACTCCC<br>GGTTTGAGGG | ATGGATCACA<br>TACCTAGTGT | ACCCCACCAT<br>TGGGGTGGTA | GAACCTTATT<br>CTTGGAATAA  | ACCGGGGTAC<br>TGGCCCCATG |

Figure 26 X

| 22751 | CCAACTCCAT<br>GGTTGAGGTA | GCTCAACAGT<br>CTTGTCA    | CCCCAGGTAC<br>GGGGTCCATG | AGCCCACQEM<br>TCGGGTGGGA | cecyecture<br>ceerocuresc |
|-------|--------------------------|--------------------------|--------------------------|--------------------------|---------------------------|
| 22801 |                          | TCTACAGCTT<br>AGATGTCGAA |                          |                          |                           |
| 22851 |                          | CAGATTAGGA<br>GTCTAATCCT |                          |                          |                           |
| 22901 | TGTAAAAATA<br>ACATTTTTAT | ATGTACTAGA<br>TACATGATCT |                          |                          |                           |
| 22951 | TTGTACACTC<br>AACATGTGAG | TCGGGTGATT<br>AGCCCACTAA |                          |                          |                           |
| 23001 |                          | AAGGGGTTCT<br>TTCCCCAAGA |                          |                          |                           |
| 23051 |                          | ATACTGGTGT<br>TATGACCACA |                          |                          |                           |
| 23101 |                          | GCTCGGTGAA<br>CGAGCCACTT |                          |                          |                           |
| 23151 | CAACGCGTTT<br>GTTGCGCAAA | AGCAGGTCGG<br>TCGTCCAGCC |                          |                          |                           |
| 23201 |                          | CGCGCGCGAG               |                          |                          |                           |
| 23251 |                          | CCGGGTGGTG               |                          |                          |                           |
| 23301 |                          | TCCAGGTCCT<br>AGGTCCAGGA |                          |                          |                           |
| 23351 |                          | CCTTCCCAAA<br>GGAAGGGTTT |                          |                          |                           |
| 23401 | TCGCACCGTA<br>AGCGTGGCAT | GTGGCATCAA<br>CACCGTAGTT |                          |                          |                           |
| 23451 | ATACAGCGCC<br>TATGTCGCGG |                          |                          |                          | ACCTGAGCCT<br>TGGACTCGGA  |
| 23501 | TTGCGCCTTC<br>AACGCGGAAG |                          |                          |                          | AAACTGATTG<br>TTTGACTAAC  |
|       |                          | GGCGCAGCAC               | GTGCGTCGTG               | GAACGCAGCC               | ACAACCTCTA                |
| 23601 | CTGCACCACA<br>GACGTGGTGT |                          |                          |                          | GCCTTGCTAG<br>CGGAACGATC  |
| 23651 |                          |                          |                          |                          | ATCCATTTCA<br>TAGGTAAAGT  |

Figure 26 Y

| 23701  | ATCACGTGCT               | CC PATTTAT               | CATAATGCTT               | CCGTGTAGAC                | ACTTAA TC                 |
|--------|--------------------------|--------------------------|--------------------------|---------------------------|---------------------------|
|        | TAGTGCACGA               | GGAATAAATA               | GTATTACGAA               | GGCACATCTG                | TGAATTCGAG                |
| 23751  |                          | TCAGCGCAGC<br>AGTCGCGTCG |                          |                           |                           |
| 23801  | CGTGATGCTT               | GTAGGTCACC               | TCTGCAAACG               | ACTGCAGGTA                | CGCCTGCAGG                |
|        | GCACTACGAA               | CATCCAGTGG               | AGACGTTTGC               | TGACGTCCAT                | GCGGACGTCC                |
| 23851  | AATCGCCCA                | TCATCGTCAC               | AAAGGTCTTG               | TTGCTGGTGA                | AGGTCAGCTG                |
|        | TTAGCGGGGT               | AGTAGCAGTG               | TTTCCAGAAC               | AACGACCACT                | TCCAGTCGAC                |
| 23901  | CAACCCGCGG               | TGCTCCTCGT               | TCAGCCAGGT               | CTTGCATACG                | GCCGCCAGAG                |
|        | GTTGGGCGCC               | ACGAGGAGCA               | AGTCGGTCCA               | GAACGTATGC                | CGGCGGTCTC                |
| 23951  | CTTCCACTTG               | GTCAGGCAGT               | AGTTTGAAGT               | TCGCCTTTAG                | ATCGTTATCC                |
|        | GAAGGTGAAC               | CAGTCCGTCA               | TCAAACTTCA               | AGCGGAAATC                | TAGCAATAGG                |
| 24001  | ACGTGGTACT               | TGTCCATCAG               | CGCGCGCGCA               | GCCTCCATGC                | CCTTCTCCCA                |
|        | TGCACCATGA               | ACAGGTAGTC               | GCGCGCGCGT               | CGGAGGTACG                | GGAAGAGGGT                |
| 24,051 | CGCAGACACG               | ATCGGCACAC               | TCAGCGGGTT               | CATCACCGTA                | ATTTCACTTT                |
|        | GCGTCTGTGC               | TAGCCGTGTG               | AGTCGCCCAA               | GTAGTGGCAT                | TAAAGTGAAA                |
| 24101  | CCGCTTCGCT               | GGGCTCTTCC               | TCTTCCTCTT               | GCGTCCGCAT                | ACCACGCGCC                |
|        | GGCGAAGCGA               | CCCGAGAAGG               | AGAAGGAGAA               | CGCAGGCGTA                | TGGTGCGCGG                |
| 24151  | ACTGGGTCGT               | CTTCATTCAG               | CCGCCGCACT               | GTGCGCTTAC                | CTCCTTTGCC                |
|        | TGACCCAGCA               | GAAGTAAGTC               | GGCGGCGTGA               | CACGCGAATG                | GAGGAAACGG                |
| 24201  | ATGCTTGATT               | AGCACCGGTG               | GGTTGCTGAA               | ACCCACCATT                | TGTAGCGCCA                |
|        | TACGAACTAA               | TCGTGGCCAC               | CCAACGACTT               | TGGGTGGTAA                | ACATCGCGGT                |
| 24251  | CATCTTCTCT<br>GTAGAAGAGA | TTCTTCCTCG<br>AAGAAGGAGC | CTGTCCACGA<br>GACAGGTGCT | TTACCTCTGG AATGGAGACC     | TGATGGCGGG<br>ACTACCGCCC  |
| 24301  | CGCTCGGGCT<br>GCGAGCCCGA | TGGGAGAAGG<br>ACCTCTTCC  | GCGCTTCTTT               | TTCTTCTTGG<br>AAGAAGAACC  | GCGCAATGGC<br>CGCGTTACCG  |
| 24351  | CAAATCCGCC<br>GTTTAGGCGG | GCCGAGGTCG               | ATGGCCGCGG<br>TACCGGCGCC | GCTGGGTGTG<br>CGACCCACAC  | CGCGGCACCA<br>CGCGCCGTGGT |
| 24401  | GCGCGTCTTG               | TGATGAGTCT               | TCCTCGTCCI               | CGGACTCGAT                | ACGCCGCCTC                |
|        | CGCGCAGAAC               | ACTACTCAGA               | AGGAGCAGGA               | CCCTGAGCTA                | TGCGGCGGAG                |
| 24451  | ATCCGCTTTT<br>TAGGCGAAAA | TTGGGGGCGC               | CCGGGGAGGC<br>GGCCCTCCG  | CCGCCGCTGC                | GGGACGGGGA<br>CCCTGCCCCT  |
| 24501  | CGACACGTCC               | TCCATGGTTG               | GGGGACGTCG               | CGCCGCACCG                | CGTCCGCGCT                |
|        | GCTGTGCAGG               | AGGTACCAAC               | CCCCTGCAGC               | CGCGCGTGGC                | CGCAGGCGCGA               |
| 24551  | CGGGGGTGGT               | TTCGCGCTGC               | TCCTCTTCCC<br>AGGAGAAGGG | GACTGGCCAT<br>G CTGACCGGT | TTCCTTCTCC<br>AAGGAAGAGG  |
| 24601  | TATAGGCAGA<br>ATATCCGTCT | AAAAGATCAT               | GGAGTCAGTCAGTCAG         | GAGAAGAAG(<br>CTCTTCTTC(  | ACAGCCTAAC<br>TGTCGGATTG  |

Figure 262

| 24651 | CCCCCCTCT<br>GCGGGGGAGA  |             |   | CACCGATGCC<br>GTGGCTACGG |                          |
|-------|--------------------------|-------------|---|--------------------------|--------------------------|
| 24701 | CTACCACCTT<br>GATGGTGGAA |             |   | TTGAGGAGGA<br>AACTCCTCCT |                          |
| 24751 | ATCGAGCAGG<br>TAGCTCGTCC | -           |   | GACGACGAGG<br>CTGCTGCTCC |                          |
| 24801 |                          |             |   | CAACGCAGAG<br>GTTGCGTCTC |                          |
| 24851 |                          | <del></del> |   | GCGACTACCT<br>CGCTGATGGA |                          |
| 24901 |                          |             |   | CAGTGCGCCA<br>GTCACGCGGT |                          |
| 24951 |                          |             |   | CGCCATAGCG<br>GCGGTATCGC | •                        |
| 25001 |                          |             |   | GCGTACCCCC<br>CGCATGGGGG | =                        |
| 25051 |                          |             | • | CTCAACTTCT<br>GAGTTGAAGA | - · · · ·                |
| 25101 | TGCCGTGCCA<br>ACGGCACGGT |             |   | CATCTTTTTC<br>GTAGAAAAAG |                          |
| 25151 |                          |             |   | GCCGAGCGGA<br>CGGCTCGCCT |                          |
| 25201 |                          |             |   | ATCGCCTCGC<br>TAGCGGAGCG |                          |
| 25251 |                          |             | - | CGAGAAGCGC<br>GCTCTTCGCG | -                        |
| 25301 |                          |             |   | GTCACTCTGG<br>CAGTGAGACC |                          |
| 25351 |                          |             |   | GTACTAAAAC<br>CATGATTTTG |                          |
| 25401 |                          |             |   |                          | AAGGTCATGA<br>TTCCAGTACT |
| 25451 |                          |             |   | GTGCGCAGCC<br>CACGCGTCGG | CCTGGAGAGG<br>GGACCTCTCC |
| 25501 |                          |             |   | GGCCTACCCG<br>CCGGATGGGC | CAGTTGGCGA<br>GTCAACCGCT |
| 25551 |                          |             |   |                          | GACTTGGAGG<br>CTGAACCTCC |

7 igure 26 AA

| 25601 | AGCGACGCAA               | AATGATG                  | GCCGCAGTGC                | TCGTTACCGT               | GGAGCT AG                |
|-------|--------------------------|--------------------------|---------------------------|--------------------------|--------------------------|
|       | TCGCTGCGTT               | TGATTACTAC               | CGGCGTCACG                | AGCAATGGCA               | CCTCGAACTC               |
| 25651 | TGCATGCAGC               | GGTTCTTTGC               | TGACCCGGAG                | ATGCAGCGCA               | AGCTAGAGGA               |
|       | ACGTACGTCG               | CCAAGAAACG               | ACTGGGCCTC                | TACGTCGCGT               | TCGATCTCCT               |
| 25701 | AACATTGCAC               | TACACCTTTC               | GACAGGGCTA                | CGTACGCCAG               | GCCTGCAAGA               |
|       | TTGTAACGTG               | ATGTGGAAAG               | CTGTCCCGAT                | GCATGCGGTC               | CGGACGTTCT               |
| 25751 | TCTCCAACGT               | GGAGCTCTGC               | AACCTGGTCT                | CCTACCTTGG               | AATTTTGCAC               |
|       | AGAGGTTGCA               | CCTCGAGACG               | TTGGACCAGA                | GGATGGAACC               | TTAAAACGTG               |
| 25801 | GAAAACCGCC               | TTGGGCAAAA               | CGTGCTTCAT                | TCCACGCTCA               | AGGGCGAGGC               |
|       | CTTTTGGCGG               | AACCCGTTTT               | GCACGAAGTA                | AGGTGCGAGT               | TCCCGCTCCG               |
| 25851 | GCGCCGCGAC               | TACGTCCGCG               | ACTGCGTTTA                | CTTATTTCTA               | TGCTACACCT               |
|       | CGCGGCGCTG               | ATGCAGGCGC               | TGACGCAAAT                | GAATAAAGAT               | ACGATGTGGA               |
| 25901 | GGCAGACGGC               | CATGGGCGTT               | TGGCAGCAGT                | GCTTGGAGGA               | GTGCAACCTC               |
|       | CCGTCTGCCG               | GTACCCGCAA               | ACCGTCGTCA                | CGAACCTCCT               | CACGTTGGAG               |
| 25951 | AAGGAGCTGC               | AGAAACTGCT               | AAAGCAAAAC                | TTGAAGGACC               | TATGGACGGC               |
|       | TTCCTCGACG               | TCTTTGACGA               | TTTCGTTTTG                | AACTTCCTGG               | ATACCTGCCG               |
| 26001 |                          | CGCTCCGTGG<br>GCGAGGCACC |                           |                          | ATTTTCCCCG<br>TAAAAGGGGC |
| 26051 | AACGCCTGCT               | TAAAACCCTG               | CAACAGGGTC                | TGCCAGACTT               | CACCAGTCAA               |
|       | TTGCGGACGA               | ATTTTGGGAC               | GTTGTCCCAG                | ACGGTCTGAA               | GTGGTCAGTT               |
| 26101 |                          |                          |                           |                          | CAGGAATCTT<br>GTCCTTAGAA |
| 26151 | GCCCGCCACC<br>CGGGCGGTGG | TGCTGTGCAC<br>ACGACACGTG | TTCCTAGCGA<br>AAGGATCGCT  | CTTTGTGCCC               | ATTAAGTACC<br>TAATTCATGG |
| 26201 | GCGAATGCCC               | TCCGCCGCTT               | TGGGGCCACT                | GCTACCTTCT               | GCAGCTAGCC               |
|       | CGCTTACGGG               | AGGCGGCGAA               | ACCCCGGTGA                | CGATGGAAGA               | CGTCGATCGG               |
| 26251 |                          |                          |                           |                          | GCGGTGACGG<br>CGCCACTGCC |
| 26301 | TCTACTGGAG<br>AGATGACCTC | TGTCACTGTC               | GCTGCAACCT<br>CGACGTTGGA  | ATGCACCCCG               | CACCGCTCCC<br>CTGGCGAGGG |
| 26351 | TGGTTTGCAA               | TTCGCAGCTG               | CTTAACGAAA                | GTCAAATTAI               | CGGTACCTTT               |
|       | ACCAAACGTT               | AAGCGTCGAC               | GAATTGCTTT                | CAGTTTAATA               | CCCATGGAAA               |
| 26401 | GAGCTGCAGG               | GTCCCTCGCC               | TGACGAAAAG                | TCCGCGGCTC               | CGGGGTTGAA               |
|       | CTCGACGTCC               | CAGGGAGCGG               | ACTGCTTTTC                | AGGCGCCGAG               | GCCCCAACTT               |
| 26451 | ACTCACTCCG<br>TGAGTGAGGC | GGGCTGTGGA               | CGTCGGCTTA<br>CGCAGCCGAAT | CCTTCGCAAA<br>GGAAGCGTTT | TTTGTACCTG AAACATGGAC    |
| 26501 | AGGACTACCA               | CGCCCACGAG               | ATTAGGTTCT                | ACGAAGACCA               | ATCCCGCCCG               |
|       | TCCTGATGGT               | CGCGGGTGCTC              | TAATCCAAGA                | TGCTTCTGG1               | TAGGGCGGGC               |

Figure 26 AB

| 26551 |   |       | ACCCAGGGCC<br>TGGGTCCCGG |                          |
|-------|---|-------|--------------------------|--------------------------|
| 26601 |   | <br>  | AGAGTTTCTG<br>TCTCAAAGAC |                          |
| 26651 |   |       | GCGAGGAGCT<br>CGCTCCTCGA |                          |
| 26701 |   | <br>  | Gececcceee               |                          |
| 26751 |   |       | CGCCGCCACC<br>GCGGCGGTGG |                          |
| 26801 | • | <br>  | GTTTTGGACG<br>CAAAACCTGC |                          |
| 26851 | • | <br>  | CGAGGAAGCT<br>GCTCCTTCGA |                          |
| 26901 |   |       | CGGTCGCATT<br>GCCAGCGTAA |                          |
| 26951 |   |       | ATGGCTACAA<br>TACCGATGTT |                          |
| 27001 |   |       | ACCCAACCGT<br>TGGGTTGGCA |                          |
| 27051 |   |       | AGCCGCCGCC<br>TCGGCGGCGG |                          |
| 27101 |   | <br>• | TGGCGCGGGC<br>ACCGCGCCCG |                          |
| 27151 | - |       | CAACATCTCC<br>GTTGTAGAGG |                          |
| 27201 |   | <br>  | TCCCCCGTAA<br>AGGGGGCATT | CATCCTGCAT<br>GTAGGACGTA |
| 27251 |   |       |                          | GCGGCAGCAA<br>CGCCGTCGTT |
| 27301 |   |       |                          | GACTCTGACA<br>CTGAGACTGT |
| 27351 |   |       |                          | GAGCGCTGCG<br>CTCGCGACGC |
| 27401 |   |       |                          | AACAGGATTT<br>TTGTCCTAAA |
| 27451 |   |       |                          | AGAACAAGAG<br>TCTTGTTCTC |

Figure 26 AC

| 27501 | CTGAAAATAA<br>GACTTTTATT | ACAGGTC<br>TITTGTCCAG    | TCTGCGATCC<br>AGACGCTAGG | CTCACCCGCA<br>GAGTGGGCGT | GCTGCC 'A<br>CGACGGACAT  |
|-------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| 27551 |                          | GAAGATCAGC<br>CTTCTAGTCG |                          |                          |                          |
| 27601 |                          | ATACTGCGCG<br>TATGACGCGC |                          |                          |                          |
| 27651 |                          | AAGCGCGAAA<br>TTCGCGCTTT |                          |                          |                          |
| 27701 |                          | GTTGTCAGCG<br>CAACAGTCGC |                          |                          |                          |
| 27751 |                          | TTACCAGCCA<br>AATGGTCGGT |                          |                          |                          |
| 27801 |                          | CCCGAATAAA<br>GGGCTTATTT |                          |                          |                          |
| 27851 |                          | GGAATACGCG<br>CCTTATGCGC |                          |                          |                          |
| 27901 |                          | CACCACACCT<br>GTGGTGTGGA |                          |                          |                          |
| 27951 |                          | TGTACCAGGA<br>ACATGGTCCT |                          |                          | TGGTACTTCC<br>ACCATGAAGG |
| 28001 |                          | CAGGCCGAAG<br>GTCCGGCTTC |                          |                          | GCGCAGCTTG<br>CGCGTCGAAC |
| 28051 |                          |                          |                          |                          | TATAACTCAC<br>ATATTGAGTG |
| 28101 |                          |                          |                          |                          | CGGTGAGCTC<br>GCCACTCGAG |
| 28151 |                          |                          |                          |                          | CCCCGCCCGCC              |
| 28201 | GCTCTTCATT<br>CGAGAAGTAA | CACGCCTCGT<br>GTGCGGAGCA | CAGGCAATCC<br>GTCCGTTAGG | TAACTCTGCA<br>ATTGAGACGT | GACCTCGTCC<br>CTGGAGCAGG |
| 28251 | TCTGAGCCGC<br>AGACTCGGCG | GCTCTGGAGG<br>CGAGACCTCC | CATTGGAACT<br>GTAACCTTGA | CTGCAATTTA<br>GACGTTAAAT | TTGAGGAGTT AACTCCTCAA    |
| 28301 | TGTGCCATCG<br>ACACGGTAGC | GTCTACTTTA<br>CAGATGAAAT | ACCCCTTCTC<br>TGGGGAAGAG | GGGACCTCCC<br>CCCTGGAGGG | GGCCACTATC<br>CCGGTGATAG |
| 28351 | CGGATCAATT<br>GCCTAGTTAA | TATTCCTAAC<br>ATAAGGATTG | TTTGACGCGGG              | TAAAGGACTC<br>ATTTCCTGAG | GGCGGACGGC               |
| 28401 | TACGACTGAA<br>ATGCTGACTT | TGTTAAGTGG<br>ACAATTCACC | AGAGGCAGAG<br>TCTCCGTCTC | CAACTGCGCG               | TGAAACACCT<br>ACTTTGTGGA |

Figure 26 AD

| 28451 |            |            |            | CCGCGACTCC               |                          |
|-------|------------|------------|------------|--------------------------|--------------------------|
|       | CCAGGTGACA | GCGGCGTGT  | TCACGAAACG | GGCGCTGAGG               | CCACTCAAAA               |
| 28501 |            | _          |            | AGGGCCCGGC<br>TCCCGGGCCG |                          |
| 28551 | CGGCTTACCG | CCCAGGGAGA | GCTTGCCCGT | AGCCTGATTC               | GGGAGTTTAC               |
|       | GCCGAATGGC | GGGTCCCTCT | CGAACGGGCA | TCGGACTAAG               | CCCTCAAATG               |
| 28601 |            |            |            | GGGACCCTGT<br>CCCTGGGACA |                          |
| 28651 |            |            |            | ATCAAGATCT<br>TAGTTCTAGA |                          |
| 28701 |            |            |            | TAAAATATAC<br>ATTTTATATG |                          |
| 28751 |            |            |            | CCCGCCCAAG<br>GGGCGGGTTC |                          |
| 28801 |            | • - •      |            | CCCTCTGTGA<br>GGGAGACACT |                          |
| 28851 |            |            |            | GAACCTCTCC<br>CTTGGAGAGG |                          |
| 28901 |            | -          |            | CCTGCCGGGA<br>GGACGGCCCT |                          |
| 28951 |            |            |            | CCTGACCGTA<br>GGACTGGCAT | •                        |
| 29001 | -          |            | -          | CCAGAACAGG<br>GGTCTTGTCC |                          |
| 29051 |            |            |            | GCAGCTACTG<br>CGTCGATGAC |                          |
| 29101 |            |            |            | TAATTCAGGT<br>ATTAAGTCCA |                          |
| 29151 |            |            |            | TTCTCTTTAT<br>AAGAGAAATA | TCTTATACTA<br>AGAATATGAT |
| 29201 |            |            |            |                          | TTTGCATTTA<br>AAACGTAAAT |
| 29251 |            |            |            |                          | TTAGGTACAT<br>AATCCATGTA |
| 29301 |            |            |            |                          | ACCCAAAAGG<br>TGGGTTTTCC |
| 29351 |            |            |            |                          | TGAAGCTAAT<br>ACTTCGATTA |

Figure 26 AE

| 29401 | CTCACGTGGT               | GAGAATATTT | ATGCACCACA<br>TACGTGGTGT | CTTGTACTTT               | TCGACGAATA               |
|-------|--------------------------|------------|--------------------------|--------------------------|--------------------------|
| 29451 | TCGCCACAAA               | AACAAAATTG | GCAAGTATGC               | TGTTTATGCT               | ATTTGGCAGC               |
|       | AGCGGTGTTT               | TTGTTTTAAC | CGTTCATACG               | ACAAATACGA               | TAAACCGTCG               |
| 29501 | CAGGTGACAC               | TACAGAGTAT | AATGTTACAG               | TTTTCCAGGG               | TAAAAGTCAT               |
|       | GTCCACTGTG               | ATGTCTCATA | TTACAATGTC               | AAAAGGTCCC               | ATTTCAGTA                |
| 29551 | AAAACTTTTA               | TGTATACTTT | TCCATTTTAT               | GAAATGTGCG               | ACATTACCAT               |
|       | TTTTGAAAAT               | ACATATGAAA | AGGTAAAATA               | CTTTACACGC               | TGTAATGGTA               |
| 29601 | GTACATGAGC               | AAACAGTATA | AGTTGTGGCC               | CCCACAAAAT               | TGTGTGGAAA               |
|       | CATGTACTCG               | TTTGTCATAT | TCAACACCGG               | GGGTGTTTTA               | ACACACCTTT               |
| 29651 |                          |            | ACTGCTATGC<br>TGACGATACG |                          |                          |
| 29701 | GTCTGTACCC               | TACTCTATAT | TAAATACAAA               | AGCAGACGCA               | GCTTTATTGA               |
|       | CAGACATGGG               | ATGAGATATA | ATTTATGTTT               | TCGTCTGCGT               | CGAAATAACT               |
| 29751 | GGAAAAGAAA               | ATGCCTTAAT | TTACTAAGTT               | ACAAAGCTAA               | TGTCACCACT               |
|       | CCTTTTCTTT               | TACGGAATTA | AATGATTCAA               | TGTTTCGATT               | ACAGTGGTGA               |
| 29801 | AACTGCTTTA               | CTCGCTGCTT | GCAAAACAAA               | TTCAAAAAGT               | TAGCATTATA               |
| !     | TTGACGAAAT               | GAGCGACGAA | CGTTTTGTTT               | AAGTTTTTCA               | ATCGTAATAT               |
| 29851 | ATTAGAATAG               | GATTTAAACC | CCCCGGTCAT               | TTCCTGCTCA               | ATACCATTCC               |
|       | TAATCTTATC               | CTAAATTTGG | GGGGCCAGTA               | AAGGACGAGT               | TATGGTAAGG               |
| 29901 | CCTGAACAAT               | TGACTCTATG | TGGGATATGC               | TCCAGCGCTA               | CAACCTTGAA               |
|       | GGACTTGTTA               | ACTGAGATAC | ACCCTATACG               | AGGTCGCGAT               | GTTGGAACTT               |
| 29951 | GTCAGGCTTC               | CTGGATGTCA | GCATCTGACT               | TTGGCCAGCA               | CCTGTCCCGC               |
|       | CAGTCCGAAG               | GACCTACAGT | CGTAGACTGA               | AACCGGTCGT               | GGACAGGGCG               |
| 30001 | GGATTTGTTC               | CAGTCCAACT | ACAGCGACCC               | ACCCTAACAG               | AGATGACCAA               |
|       | CCTAAACAAG               | GTCAGGTTGA | TGTCGCTGGG               | TGGGATTGTC               | TCTACTGGTT               |
| 30051 | CACAACCAAC<br>GTGTTGGTTG | cecceccec  | CTACCGGACT<br>GATGGCCTGA | TACATCTACC<br>ATGTAGATGG | ACAAATACAC<br>TGTTTATGTG |
| 30101 | CCCAAGTTTC               | TGCCTTTGTC | AATAACTGGG               | ATAACTTGGG               | CATGTGGTGG               |
|       | GGGTTCAAAG               | ACGGAAACAG | TTATTGACCC               | TATTGAACCC               | GTACACCACC               |
| 30151 | TTCTCCATAG<br>AAGAGGTATC | CGCTTATGTT | TGTATGCCTT<br>ACATACGGAA | ATTATTATGT<br>TAATAATACA | GGCTCATCTG<br>CCGAGTAGAC |
| 30201 | CTGCCTAAAG<br>GACGGATTTC | CGCAAACGCG | CCCGACCACC<br>GGGCTGGTGG | CATCTATAGT<br>GTAGATATCA | CCCATCATTG<br>GGGTAGTAAC |
| 30251 | TGCTACACCC               | AAACAATGAT | GGAATCCATA               | GATTGGACGG               | ACTGAAACAC               |
|       | ACGATGTGGG               | TTTGTTACTA | CCTTAGGTAT               | CTAACCTGCC               | TGACTTTGTG               |
| 30301 | ATGTTCTTTT               | CTCTTACAGI | ATGATTAAAT               | GAGACATGAT               | TCCTCGAGTT               |
|       | TACAAGAAAA               | GAGAATGTCA | TACTAATTTA               | CTCTGTACTA               | AGGAGCTCAA               |

Figure 26 AF

| 30351 |                          | T CCTTGT<br>ACTGGGAACA   |            |            | CCACAT C.                |
|-------|--------------------------|--------------------------|------------|------------|--------------------------|
| 30401 | TGCGGTTTCT<br>ACGCCAAAGA | CACATCGAAG<br>GTGTAGCTTC |            |            |                          |
| 30451 |                          | ATTTGTCACC<br>TAAACAGTGG |            |            |                          |
| 30501 |                          | TTATCCAGTG<br>AATAGGTCAC |            |            |                          |
| 30551 |                          | CATCCCCAGT<br>GTAGGGGTCA |            |            |                          |
| 30601 |                          | ATTATGAAAT<br>TAATACTTTA |            |            |                          |
| 30651 |                          | GTTTTGTTCC<br>CAAAACAAGG |            |            | ACATATATCA<br>TGTATATAGT |
| 30701 |                          | CTCGTATATG<br>GAGCATATAC |            |            |                          |
| 30751 | •••••                    | GAAGCCTGGT<br>CTTCGGACCA |            |            |                          |
| 30801 |                          | CTTAGCCCTA<br>GAATCGGGAT |            |            |                          |
| 30851 |                          | ATGCCATGAA<br>TACGGTACTT |            |            |                          |
| 30901 |                          | CAAGTTGTTG<br>GTTCAACAAC |            |            |                          |
| 30951 |                          | TCCCACCCCC<br>AGGGTGGGGG |            |            |                          |
| 31001 |                          | GACACCCTAG<br>CTGTGGGATC | =          |            |                          |
| 31051 |                          |                          |            |            | GCGCATGAAT<br>CGCGTACTTA |
|       | GTTCTCGAGG               | TTCTGTACCA               | ATTGAACGTG | GTCACGTTTT | GGGGTATCTT<br>CCCCATAGAA |
| 31151 |                          |                          |            |            | ACCACCGGAC<br>TGGTGGCCTG |
| 31201 |                          |                          |            |            | GGTGGTCATG<br>CCACCAGTAC |
| 31251 |                          |                          |            |            | AAACCGAAGG<br>TTTGGCTTCC |

Figure 26 AG

| 31301 | CTGCATTCAC               | TCTTGTC                  | AAGGACCTGA               | GGATCTCTGC               | ACCCTT TA                |
|-------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
|       | GACGTAAGTG               | AGTGGAACAG               | TTCCTGGACT               | CCTAGAGACG               | TGGGAALAAT               |
| 31351 | AGACCCTGTG               | CGGTCTCAAA               | GATCTTATTC               | CCTTTAACTA               | AAAAAAATA                |
|       | TCTGGGACAC               | GCCAGAGTTT               | CTAGAATAAG               | GGAAATTGAT               | TTTTTTTAT                |
| 31401 |                          | ATCACTTACT<br>TAGTGAATGA |                          |                          |                          |
| 31451 | TATTCAGCAG               | CACCTCCTTG               | CCCTCCTCCC               | AGCTCTGGTA               | TTGCAGCTTC               |
|       | ATAAGTCGTC               | GTGGAGGAAC               | GGGAGGAGGG               | TCGAGACCAT               | AACGTCGAAG               |
| 31501 | CTCCTGGCTG               | CAAACTTTCT               | CCACAATCTA               | AATGGAATGT               | CAGTTTCCTC               |
|       | GAGGACCGAC               | GTTTGAAAGA               | GGTGTTAGAT               | TTACCTTACA               | GTCAAAGGAG               |
| 31551 | CTGTTCCTGT               | CCATCCGCAC               | CCACTATCTT               | CATGTTGTTG               | CAGATGAAGC               |
|       | GACAAGGACA               | GGTAGGCGTG               | GGTGATAGAA               | GTACAACAAC               | GTCTACTTCG               |
| 31601 | GCGCAAGACC               | GTCTGAAGAT               | ACCTTCAACC               | CCGTGTATCC               | ATATGACACG               |
|       | CGCGTTCTGG               | CAGACTTCTA               | TGGAAGTTGG               | GGCACATAGG               | TATACTGTGC               |
| 31651 |                          | CTCCAACTGT<br>GAGGTTGACA |                          |                          |                          |
| 31701 | CAATGGGTTT               | CAAGAGAGTC               | CCCCTGGGGT               | ACTCTCTTTG               | CGCCTATCCG               |
|       | GTTACCCAAA               | GTTCTCTCAG               | GGGGACCCCA               | TGAGAGAAAC               | GCGGATAGGC               |
| 31751 |                          | TACCTCCAAT<br>ATGGAGGTTA |                          |                          |                          |
| 31801 | CTCTCTCTGG               | ACGAGGCCGG               | CAACCTTACC               | TCCCAAAATG               | TAACCACTGT               |
|       | GAGAGAGACC               | TGCTCCGGCC               | GTTGGAATGG               | AGGGTTTTAC               | ATTGGTGACA               |
| 31851 | GAGCCCACCT               | CTCAAAAAAA               | CCAAGTCAAA               | CATAAACCTG               | GAAATATCTG               |
|       | CTCGGGTGGA               | GAGTTTTTTT               | GGTTCAGTTT               | GTATTTGGAC               | CTTTATAGAC               |
| 31901 | CACCCTCAC<br>GTGGGGAGTG  | AGTTACCTCA<br>TCAATGGAGT | GAAGCCCTAA<br>CTTCGGGATT | CTGTGGCTGC               | CGCCGCACCT               |
| 31951 | CTAATGGTCG<br>GATTACCAGC | CGGGCAACAC               | ACTCACCATG<br>TGAGTGGTAC | CAATCACAGG               | CCCCGCTAAC<br>GGGGCGATTG |
| 32001 | CGTGCACGAC<br>GCACGTGCTG | TCCAAACTTA<br>AGGTTTGAAT | GCATTGCCÀC<br>CGTAACGGTG | CCAAGGACCC               | CTCACAGTGT<br>GAGTGTCACA |
| 32051 | CAGAAGGAAA<br>GTCTTCCTTT | GCTAGCCCTG<br>CGATCGGGAC | CAAACATCAG<br>GTTTGTAGTC | GCCCCTCAC                | CACCACCGAT<br>GTGGTGGCTA |
| 32101 | AGCAGTACCO<br>TCGTCATGGG | TTACTATCAC<br>AATGATAGTG | TGCCTCACCC               | CCTCTAACTA<br>GGAGATTGAT | CTGCCACTGG<br>CACGGTGACC |
| 32151 | TAGCTTGGGC               | ATTGACTTGA               | AAGAGCCCAT               | TTATACACAA               | AATGGAAAAC               |
|       | ATCGAACCCG               | TAACTGAACT               | TTCTCGGGTA               | AATATGTGTT               | TTACCTTTTG               |
| 32201 | TAGGACTAAA               | GTACGGGGCT               | CCTTTGCATG               | TAACAGACGA               | CCTAAACACT               |
|       | ATCCTGATTT               | CATGCCCCGA               | GGAAACGTAC               | ATTGTCTGCT               | GGATTTGTGA               |

Figure 26 AH

| 32251  | TTGACCGTAG<br>AACTGGCATC | CTGGTCC | AGGTGTGACT<br>TCCACACTGA | ATTAATAATTA<br>TATTATTAAT | CTTCCT CA<br>GAAGGA GT   |
|--------|--------------------------|---------|--------------------------|---------------------------|--------------------------|
| 32301  |                          |         |                          | TTCACAAGGC<br>AAGTGTTCCG  |                          |
| 32351  | TTAATGTAGC<br>AATTACATCG |         |                          |                           |                          |
| 32401  |                          |         |                          | AACCAACTAA<br>TTGGTTGATT  |                          |
| 32451  |                          |         |                          | CCACAACTTG<br>GGTGTTGAAC  |                          |
| 32501  |                          |         |                          | CAAACAATTC<br>GTTTGTTAAG  |                          |
| 32551  |                          |         |                          | ATGTTTGACG<br>TACAAACTGC  |                          |
| 32601  |                          |         |                          | TGGTTCACCT<br>ACCAAGTGGA  |                          |
| 32651  |                          |         |                          | ATGGCCTAGA<br>TACCGGATCT  |                          |
| 32701  |                          |         |                          | GGCCTTAGTT<br>CCGGAATCAA  |                          |
| 3,2751 |                          |         |                          | TGATAAGCTA<br>ACTATTCGAT  |                          |
| 32801  |                          |         |                          | TAAATGCAGA<br>ATTTACGTCT  |                          |
| 32851  |                          |         |                          | AGTCAAATAC<br>TCAGTTTATG  |                          |
| 32901  |                          |         |                          | TCCAATATCT<br>AGGTTATAGA  |                          |
| 32951  |                          |         |                          | AAAATGGAGT<br>TTTTACCTCA  |                          |
| 33001  |                          |         |                          |                           | GAGATCTTAC<br>CTCTAGAATG |
| 33051  |                          |         |                          |                           | AACCTATCAG<br>TTGGATAGTC |
| 33101  |                          |         |                          |                           | TGTCAGTCAA<br>ACAGTCAGTT |
| 33151  |                          |         |                          |                           | CCATTACACT<br>GGTAATGTGA |

Figure 26 AI

| 33201 | AAACGGTACA    | GAAACAG                  | GAGACACAAC                     | TCCAAGTGCA               | TACTOT                   |
|-------|---------------|--------------------------|--------------------------------|--------------------------|--------------------------|
|       | TTTGCCATGT    | GLECTTTGTC               | CTCTGTGTTG                     | AGGTTCACGT               | ATGAGACA                 |
| 33251 |               |                          |                                | ACATTAATGA               |                          |
|       | GTAAAAGTAC    | CCTGACCAGA               | CCGGTGTTGA                     | TGTAATTACT               | TIMIMAACGG               |
| 33301 |               |                          |                                | CAAGAATAAA               |                          |
|       | TGTAGGAGAA    | TGTGAAAAAG               | TATGTAACGG                     | GTTCTTATTT               | CTTAGCAAAC               |
| 33351 |               |                          |                                | TGCAGAAAAT               |                          |
|       | ACAATACAAA    | GTTGCACAAA               | TAAAAAGTTA                     | ACGTCTTTTA               | AAGTTCAGTA               |
|       |               | •                        |                                |                          |                          |
| 33401 |               |                          |                                | CATAGCTTAT<br>GTATCGAATA |                          |
|       | AAAAGTAAGT    | CATCATATCG               | GGGIGGIGGI                     | GIAICGAAIA               | 1010170100               |
| 33451 | -             |                          |                                | ATTCAACCTG               |                          |
|       | CATGGAATTA    | GTTTGAGTGT               | CTTGGGATCA                     | TAAGTTGGAC               | GGTGGAGGGA               |
| 33501 |               |                          |                                | CCCGGCTGGC               |                          |
|       | GGGTTGTGTG    | TCTCATGTGT               | CAGGAAAGAG                     | GGGCCGACCG               | GAATTTTTCG               |
| 33551 | ATCATATCAT    | GGGTAACAGA               | CATATTCTTA                     | GGTGTTATAT               | TCCACACGGT               |
|       | TAGTATAGTA    | CCCATTGTCT               | GTATAAGAAT                     | CCACAATATA               | AGGTGTGCCA               |
| 33601 | TTCCTGTCGA    | GCCAAACGCT               | CATCAGTGAT                     | ATTAATAAAC               | TCCCCGGGCA               |
|       | AAGGACAGCT    | CGGTTTGCGA               | GTAGTCACTA                     | TAATTATTTG               | AGGGGCCCGT               |
| 33651 | GCTCACTTAA    | GTTCATGTCG               | CTGTCCAGCT                     | GCTGAGCCAC               | AGGCTGCTGT               |
|       | CGAGTGAATT    | CAAGTACAGC               | GACAGGTCGA                     | CGACTCGGTG               | TCCGACGACA               |
| 33701 | CCAACTTGCG    | GTTGCTTAAC               | GGGCGGCGAA                     | GGAGAAGTCC               | ACGCCTACAT               |
|       | GGTTGAACGC    | CAACGAATTG               | CCCGCCGCTT                     | CCTCTTCAGG               | TGCGGATGTA               |
| 33751 | GGGGGTAGAG    | TCATAATCGT               | GCATCAGGAT                     | AGGGCGGTGG               | TGCTGCAGCA               |
| 33.31 |               |                          |                                | TCCCGCCACC               |                          |
| 33801 | CCCCCCA AT    | ል <b>ል ል ር ጥር ር ጥር</b> ር | CGCCGCCGCT                     | CCGTCCTGCA               | GGAATACAAC               |
| 33601 | *             |                          |                                | GGCAGGACGT               |                          |
| 33851 | > DOCC > CDCC | <b>Managemen</b>         | <i>ር እ መ</i> ሮ <i>እ መ</i> ጥሮርር | ACCGCCCGCA               | CCATARCCC                |
| 33631 |               |                          |                                | TGGCGGGCGT               |                          |
|       |               |                          | > 0000 > 0000                  | 73 mama 3 amm            | ********                 |
| 33901 | CCTTGTCCTC    | CGGGCACAGC               | TCGCGTGGGA                     | GATCTCACTT<br>CTAGAGTGAA | TTTAGTCGTG               |
|       |               |                          |                                |                          |                          |
| 33951 |               |                          |                                |                          | ACAGTGCAAG<br>TGTCACGTTC |
|       |               |                          |                                |                          | •                        |
| 34001 | GCGCTGTATC    | CAAAGCTCAT               | CCCCCCTCC                      | ACAGAACCCA               | CGTGGCCATC<br>GCACCGGTAG |
|       |               |                          |                                |                          |                          |
| 34051 |               |                          |                                |                          | AACACGCTGG<br>TTGTGCGACC |
|       | TATGGTGTTC    | GCGICCMICI               | WI ICHCCOC                     | 'Anannaiu;               |                          |
| 34101 | ACATAAACAT    | TACCTCTTTT               | GGCATGTTGT                     | AATTCACCAC               | CTCCCGGTAC               |
|       | TGTATTTGTA    | ATGGAGAAAA               | CCGTACAACA                     | TTAAGTGGTG               | GAGGGCCATG               |

Figure 26 AJ

| 34151 |                          |             |            |                          | TCCTAA CA  |
|-------|--------------------------|-------------|------------|--------------------------|--|
|       | GTATATTTGG               | ASSICTAATTT | GTACCGCGGT | AGGTGGTGGT               | AGGATT   |
| 34201 |                          |             |            | CTGCAGGGAA<br>GACGTCCCTT |  |
| 21254 |                          |             |            |                          |  |
| 34251 |                          |             |            | AACCATGGAT<br>TTGGTACCTA | -  |
| 34301 | GTCATGATAT               | CAATGTTGGC  | ACAACACAGG | CACACGTGCA               | TACACTTCCT   |
| 54501 |                          |             |            | GTGTGCACGT               |  |
|       |                          |             |            |                          |  |
| 34351 | CAGGATTACA               |             |            | CATATCCCAG<br>GTATAGGGTC |  |
|       |                          |             |            |                          |  |
| 34401 |                          |             |            | AGGGAAGACC<br>TCCCTTCTGG |  |
| 34451 | ርጥሮልሮርጥጥርጥ               | GCATTGTCAA  | AGTGTTACAT | TCGGGCAGCA               | GCGGATGATC   |
| 24422 | +                        |             |            | AGCCCGTCGT               |  |
| 34501 |                          |             |            | AAAAGGAGGT               |  |
|       | GAGGTCATAC               | CATCGCGCCC  | AAAGACAGAG | TTTTCCTCCA               | TCTGCTAGGG   |
| 34551 |                          |             |            | ATCGTGTTGG               |  |
|       | ATGACATGCC               | TCACGCGGCT  | CTGTTGGCTC | TAGCACAACC               | AGCATCACAG   |
| 34601 |                          |             |            | TTTCCTGAAG<br>AAAGGACTTC |  |
|       |                          |             |            |                          |  |
| 34651 |                          |             |            | GGTCTCGCCG<br>CCAGAGCGGC | and the second s |
| 34701 | <del>ጥር</del> ጥርጥርጥልርጥ   | ልርተጥርጥልርጥል  | TATCCACTCT | CTCAAAGCAT               | CCAGGCGCCC   |
| 34701 |                          |             |            | GAGTTTCGTA               |  |
| 34751 |                          |             |            |                          | GCCCTGATAA   |
|       | GGACCGAAGC               | CCAAGATACA  | TTTGAGGAAG | TACGCGGCGA               | CGGGACTATT   |
| 34801 |                          |             |            | GCCAACCTAC               |  |
|       | GTAGGTGGTG               | GCGTCTTATT  | CGGTGTGGGT | CGGTTGGATG               | TGTAAGCAAG   |
| 34851 |                          |             |            |                          | CCATGTTTTT<br>GGTACAAAAA   |
|       |                          |             |            |                          |  |
| 34901 | TTTTTTATTC<br>AAAAAATAAG |             |            |                          | TAGATAATTC   |
| 34951 | TGAACGCGCT               | CCCCTCCGGT  | GGCGTGGTCA | AACTCTACAG               | CCAAAGAACA   |
| 54551 |                          |             |            |                          | GGTTTCTTGT   |
| 35001 | GATAATGGCA               |             |            |                          |  |
|       | CTATTACCGT               | AAACATTCTA  | CAACGTGTTA | CCGAAGGTTT               | TCCGTTTGCC   |
| 35051 | CCCTCACGTC<br>GGGAGTGCAG |             |            |                          | GTGAATCTCC<br>CACTTAGAGG   |

Figure 26 AK

| 35101 | TCTATAAACA<br>AGATATTTGT | TAGCACC<br>AAGGTCGTGG    | TTCAACCATG<br>AAGTTGGTAC | CCCAAATAAT<br>GGGTTTATTA | TCTCAT G<br>AGAGTAGAGC   |
|-------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| 35151 |                          |                          |                          | CCGAATATTA<br>GGCTTATAAT |                          |
| 35201 |                          |                          |                          | CCTTCAGCCT<br>GGAAGTCGGA |                          |
| 35251 |                          |                          | -                        | AGACCTGTAT<br>TCTGGACATA |                          |
| 35301 |                          |                          |                          | CGTAGGTCCC<br>GCATCCAGGG |                          |
| 35351 |                          |                          |                          | GACCAGCGCG<br>CTGGTCGCGC |                          |
| 35401 |                          |                          |                          | TGATTATGAC<br>ACTAATACTG |                          |
| 35451 |                          |                          |                          |                          | GCATGGGCGG<br>CGTACCCGCC |
| 35501 |                          |                          |                          |                          | GCCTCGCGCA<br>CGGAGCGCGT |
| 35551 |                          |                          |                          |                          | GCAGGTAAGC<br>CGTCCATTCG |
| 35601 |                          |                          |                          |                          | ACATGTCTGC<br>TGTACAGACG |
| 35651 |                          |                          |                          |                          | TTTAAACATT<br>AAATTTGTAA |
| 35701 |                          |                          |                          |                          | ATAAGACGGA<br>TATTCTGCCT |
| 35751 | GATGCCGGTA               | CGGCCGCACT               | GGCATTTTTT               | TGACCAGTGG               | GTGATTAAAA<br>CACTAATTTT |
| 35801 | TCGTGGTGGC               | TGTCGAGGAG               | CCAGTACAGG               | CCTCAGTATT               | TGTAAGACTC<br>ACATTCTGAG |
|       | CCATTTGTGT               | AGTCCAACTA               | AGTGTAGCCA               | GTCACGATTT               | AAGCGACCGA<br>TTCGCTGGCT |
| 35901 | AATAGCCCGG<br>TTATCGGGCC | GGGAATACAT<br>CCCTTATGTA | ACCCGCAGGC               | GTAGAGACAA<br>CATCTCTGTT | CATTACAGCC<br>GTAATGTCGG |
| 35951 |                          |                          |                          |                          | CATAAACACC               |
| 36001 | TGAAAAACCC<br>ACTTTTTGGG | TCCTGCCTAG<br>AGGACGGATC | GCAAAATAGC<br>CGTTTTATCG | ACCCTCCCGC<br>TGGGAGGGCG | TCCAGAACAA<br>AGGTCTTGTT |

Figure 26 AL

| 36051 | CATACAGCGC<br>GTATGTCGCG |                          |                          | CAGTCAGCCT<br>GTCAGTCGGA |                          |
|-------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| 36101 |                          |                          |                          | ACACGGCACC<br>TGTGCCGTGG |                          |
| 36151 | GTCACAGTGT<br>CAGTGTCACA |                          |                          | GCGAGTATAT<br>CGCTCATATA |                          |
| 36201 |                          |                          |                          | AACACCCAGA<br>TTGTGGGTCT |                          |
| 36251 | CGAACCTACG<br>GCTTGGATGC |                          |                          | AACCCACAAC<br>TTGGGTGTTG |                          |
| 36301 |                          |                          |                          | TTCCCATTTT<br>AAGGGTAAAA |                          |
| 36351 |                          |                          |                          | CTAAAACCTA<br>GATTTTGGAT |                          |
| 36401 |                          |                          |                          | ACTCCACCCC<br>TGAGGTGGGG |                          |
| •     |                          |                          |                          |                          | PacI                     |
| 36451 |                          |                          |                          | ATTGATGATG               |                          |
|       | ATAACCGAAG               | TTAGGTTTTA               | TTCCATATAA               | TAACTACTAC               | AATTAATTCT               |
| 36501 |                          |                          |                          | CTTCCCCATT               |                          |
|       | TAAGCCTAGA               | CGCTGCGCTC               | CGACCTACCG               | GAAGGGGTAA               | TACTAAGAAG               |
| 36551 |                          |                          |                          | TGCAGGCCAŢ               |                          |
|       | AGCGAAGGCC               | GCCGTAGCCC               | TACGGGCGCA               | ACGTCCGGTA               | CGACAGGTCC               |
| 36601 |                          |                          |                          | CAAGGCCAGC               |                          |
|       | GTCCATCTAC               | TGCTGGTAGT               | CCCTGTCGAA               | GTTCCGGTCG               | TTTTCCGGTC               |
| 36651 | GAACCGTAAA               | AAGGCCGCGT               | TGCTGGCGTT               | TTTCCATAGG               | CTCCGCCCCC               |
|       | CTTGGCATTT               | TTCCGGCGCA               | ACGACCGCAA               | AAAGGTATCC               | GAGGCGGGGG               |
| 36701 |                          |                          |                          | GTCAGAGGTG<br>CAGTCTCCAC |                          |
| 36751 |                          |                          |                          |                          | CCCTCGTGCG<br>GGGAGCACGC |
| 36801 | CTCTCCTGTT<br>GAGAGGACAA | CCGACCCTGC<br>GGCTGGGACG | CGCTTACCGG<br>GCGAATGGCC | ATACCTGTCC<br>TATGGACAGG | GCCTTTCTCC<br>CGGAAAGAGG |
| 36851 | CTTCGGGAAG               | CGTGGCGCTT               | TCTCATAGCT               | CACGCTGTAG               | GTATCTCAGT               |
|       | GAAGCCCTTC               | GCACCGCGAA               | AGAGTATCGA               | GTGCGACATC               | CATAGAGTCA               |
| 36901 | TCGGTGTAGG<br>AGCCACATCC | TCGTTCGCTC<br>AGCAAGCGAG | CAAGCTGGGC<br>GTTCGACCCG | TGTGTGCACG<br>ACACACGTGC | AACCCCCCGT<br>TTGGGGGGCA |

Figure 26 AM

| 36951 | TCAGCCCGAC               | TGCGCCT<br>GCGACGCGGA    | TATCCGGTAA               | CTATCGTCTT               | GAGTCC CC                |
|-------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| 37001 |                          | CGACTTATCG               |                          | •                        |                          |
| 37001 | GCCATTCTGT               | GCTGAATAGC               | GGTGACCGTC               | GTCGGTGACC               | ATTGTCCTAA               |
| 37051 |                          | GGTATGTAGG<br>CCATACATCC |                          |                          |                          |
| 37101 | TAACTACGGC               | TACACTAGAA               | GGACAGTATT               | TGGTATCTGC               | GCTCTGCTGA               |
|       |                          | ATGTGATCTT               |                          |                          |                          |
| 37151 |                          | CTTCGGAAAA<br>GAAGCCTTTT |                          |                          |                          |
| 37201 | ACCACCGCTG               | GTAGCGGTGG<br>CATCGCCACC | TTTTTTTGTT               | TGCAAGCAGC               | AGATTACGCG               |
| 37251 |                          | GGATCTCAAG               |                          | •                        |                          |
| 37231 |                          | CCTAGAGTTC               |                          |                          |                          |
| 37301 | ACGCTCAGTG<br>TGCGAGTCAC | GAACGAAAAC<br>CTTGCTTTTG | TCACGTTAAG<br>AGTGCAATTC | GGATTTTGGT<br>CCTAAAACCA | CATGAGATTA<br>GTACTCTAAT |
| 37351 | 'TCAAAAAGGA              | TCTTCACCTA<br>AGAAGTGGAT | GATCCTTTTA               | AATCAATCTA               | AAGTATATAT               |
| 37401 |                          | GGTCTGACAG               |                          |                          |                          |
| 3/401 | CTCATTTGAA               | CCAGACTGTC               | AATGGTTACG               | AATTAGTCAC               | TCCGTGGATA               |
| 37451 | CTCAGCGATC<br>GAGTCGCTAG | TGTCTATTTC<br>ACAGATAAAG | GTTCATCCAT<br>CAAGTAGGTA | AGTTGCCTGA<br>TCAACGGACT | CTCCCCGTCG<br>GAGGGGCAGC |
| 37501 | TGTAGATAAC               | TACGATACGG               | GAGGGCTTAC               | CATCTGGCCC               | CAGTGCTGCA<br>GTCACGACGT |
| 37551 | -                        |                          |                          |                          | CAGCAATAAA               |
| 3/331 | TACTATGGCG               | CTCTGGGTGC               | GAGTGGCCGA               | GGTCTAAATA               | GTCGTTATTT               |
| 37601 | CCAGCCAGCC<br>GGTCGGTCGG | GGAAGGGCCG<br>CCTTCCCGGC | AGCGCAGAAG<br>TCGCGTCTTC | TGGTCCTGCA<br>ACCAGGACGT | ACTTTATCCG<br>TGAAATAGGC |
| 37651 | CCTCCATCCA               | GTCTATTAAT               | TGTTGCCGG                | AAGCTAGAGT               | AAGTAGTTCG               |
| 22201 |                          |                          |                          |                          | TTCATCAAGC<br>GCATCGTGGT |
| 3//01 | GGTCAATTAT               | CAAACGCGTT               | GCAACAACGG               | TAACGATGTC               | CGTAGCACCA               |
| 37751 | GTCACGCTCG<br>CAGTGCGAGC | TCGTTTGGTA<br>AGCAAACCAT | TGGCTTCATT<br>ACCGAAGTAA | CAGCTCCGGT<br>GTCGAGGCCA | TCCCAACGAT<br>AGGGTTGCTA |
| 37801 | CAAGGCGAGT<br>GTTCCGCTCA | TACATGATCC<br>ATGTACTAGG | CCCATGTTGT<br>GGGTACAACA | GCAAAAAAGC<br>CGTTTTTCG  | GGTTAGCTCC<br>CCAATCGAGG |
| 37851 | TTCGGTCCTC<br>AAGCCAGGAG | CGATCGTTGT               | CAGAAGTAAG<br>GTCTTCATTC | TTGGCCGCAG               | TGTTATCACT ACAATAGTGA    |

Figure 26 AN

| 37901 | CATGGTTATG<br>GTACCAATAC | CE-CGTGACG | ATAATTCTCT<br>TATTAAGAGA | TACTGTCATG<br>ATGACAGTAC | CCATCO AA * |
|-------|--------------------------|------------|--------------------------|--------------------------|-------------|
| 37951 | GATGCTTTTC               | TGTGACTGGT | GAGTACTCAA               | CCAAGTCATT               | CTGAGAATAG  |
|       | CTACGAAAAG               | ACACTGACCA | CTCATGAGTT               | GGTTCAGTAA               | GACTCTTATC  |
| 38001 | TGTATGCGGC               | GACCGAGTTG | CTCTTGCCCG               | GCGTCAACAC               | GGGATAATAC  |
|       | ACATACGCCG               | CTGGCTCAAC | GAGAACGGGC               | CGCAGTTGTG               | CCCTATTATG  |
| 38051 | CGCGCCACAT               | AGCAGAACTT | TAAAAGTGCT               | CATCATTGGA               | AAACGTTCTT  |
|       | GCGCGGTGTA               | TCGTCTTGAA | ATTTTCACGA               | GTAGTAACCT               | TTTGCAAGAA  |
| 38101 |                          | ACTCTCAAGG |                          |                          |             |
|       | GCCCCGCTTT               | TGAGAGTTCC | TAGAATGGCG               | ACAACTCTAG               | GTCAAGCTAC  |
| 38151 | TAACCCACTC               | GTGCACCCAA | CTGATCTTCA               | GCATCTTTTA               | CTTTCACCAG  |
|       | ATTGGGTGAG               | CACGTGGGTT | GACTAGAAGT               | CGTAGAAAAT               | GAAAGTGGTC  |
| 38201 |                          |            |                          |                          | AAAAAGGGAA  |
|       | GCAAAGACCC               | ACTCGTTTTT | GTCCTTCCGT               | TTTACGGCGT               | TTTTTCCCTT  |
| 38251 |                          |            |                          |                          | TTTTCAATAT  |
|       | ATTCCCGCTG               | TGCCTTTACA | ACTTATGAGT               | ATGAGAAGGA               | AAAAGTTATA  |
| 38301 |                          |            |                          |                          | ACATATTTGA  |
|       | v                        | AAATAGTCCC | AATAACAGAG               | TACTCGCCTA               | TGTATAAACT  |
| 38351 |                          |            |                          |                          | TTTCCCCGAA  |
|       | TACATAAATC               | TTTTTATTTG | TTTATCCCCA               | AGGCGCGTGT               | AAAGGGGCTT  |
| 38401 |                          |            |                          |                          | ATTAACCTAT  |
|       |                          |            |                          |                          | TAATTGGATA  |
| 38451 |                          |            |                          |                          | TGGATCCGAA  |
|       | TTTTTATCCG               | CATAGTGCTC | CGGGAAAGCA               | GAAGTTCTTA               | ACCTAGGCTT  |
|       | -                        | PacI       |                          |                          |             |
| 38501 | TTCTTAATTT               | CTTAATTAA  | (SEQ ID NO               | :32)                     |             |

AAGAATTAAA GAATTAATT (SEQ ID NO:33)

Figure 26 AO

## MRKAd5nef MER1063 (MRKAd5 Pre-Adenoviral Vector Containing the G2A,LLA nef Coding Region)

| 1   | CATCATCAAT    | AATATACCTT | ATTTTGGATT | GAAGCCAATA | TGATAATGAG |
|-----|---------------|------------|------------|------------|------------|
|     | GTAGTAGTTA    | TTATATGGAA | TAAAACCTAA | CTTCGGTTAT | ACTATTACTC |
| 51  | GGGGTGGAGT    | TTGTGACGTG | GCGCGGGGG  | TGGGAACGGG | GCGGGTGACG |
|     |               | •          |            |            | CGCCCACTGC |
| 101 | TAGTAGTGTG    | GCGGAAGTGT | GATGTTGCAA | GTGTGGCGGA | ACACATGTAA |
|     |               |            |            |            | TGTGTACATT |
| 151 | GCGACGGATG    | TGGCAAAAGT | GACGTTTTTG | GTGTGCGCCG | GTGTACACAG |
|     |               |            |            |            | CACATGTGTC |
| 201 | GAAGTGACAA    | TTTTCGCGCG | GTTTTAGGCG | GATGTTGTAG | TAAATTTGGG |
|     |               |            |            | •          | ATTTAAACCC |
| 251 | CGTAACCGAG    | TAAGATTTGG | CCATTTTCGC | GGGAAAACTG | AATAAGAGGA |
|     |               |            |            |            | TTATTCTCCT |
| 301 | AGTGAAATCT    | GAATAATTTT | GTGTTACTCA | TAGCGCGTAA | TATTTGTCTA |
|     |               |            |            |            | ATAAACAGAT |
| 351 | GGGCCGCGGG    | GACTTTGACC | GTTTACGTGG | AGACTCGCCC | AGGTGTTTTT |
|     |               |            |            |            | TCCACAAAAA |
| 401 | CTCAGGTGTT    | TTCCGCGTTC | CGGGTCAAAG | TTGGCGTTTT | ATTATTATAG |
|     |               |            |            |            | TAATAATATC |
| 451 | GCGGCCGCGA    | TCCATTGCAT | ACGTTGTATC | CATATCATAA | TATGTACATT |
|     |               |            |            |            | ATACATGTAA |
| 501 | TATATTGGCT    | CATGTCCAAC | ATTACCGCCA | TGTTGACATT | GATTATTGAC |
|     |               |            |            |            | CTAATAACTG |
| 551 | TAGTTATTAA    | TAGTAATCAA | TTACGGGGTC | ATTAGTTCAT | AGCCCATATA |
|     |               |            |            |            | TCGGGTATAT |
| 601 | TGGAGTTCCG    | CGTTACATAA | CTTACGGTAA | ATGGCCCGCC | TGGCTGACCG |
|     |               |            |            |            | ACCGACTGGC |
| 651 | CCCAACGACC    | CCCGCCCATT | GACGTCAATA | ATGACGTATG | TTCCCATAGT |
|     |               |            |            |            | AAGGGTATCA |
| 701 | AACGCCAATA    | GGGACTTTCC | ATTGACGTCA | ATGGGTGGAG | TATTTACGGT |
|     |               |            |            |            | ATAAATGCCA |
| 751 | AAACTGCCCA    | CTTGGCAGTA | CATCAAGTGT | ATCATATGCC | AAGTACGCCC |
|     |               |            |            |            | TTCATGCGGG |
| 801 | CCTATTGACG    | TCAATGACGG | TAAATGGCCC | GCCTGGCATT | ATGCCCAGTA |
|     | ርርኔ ጥል አርጥርርር | AGTTACTGCC | ATTTACCGGG | CGGACCGTAA | TACGGGTCAT |

Figure 27A

| 851          |                          |                          |                          | GTACATCTAC<br>CATGTAGATG  |                          |
|--------------|--------------------------|--------------------------|--------------------------|---------------------------|--------------------------|
| 901          |                          |                          |                          | AGTACATCAA                |                          |
|              |                          |                          |                          | TCATGTAGTT                |                          |
| 951          |                          |                          |                          | GAGGTGGGGT                |                          |
| 1001         |                          |                          |                          | GGACTTTCCA<br>CCTGAAAGGT  |                          |
| 1051         |                          |                          |                          | GTAGGCGTGT<br>CATCCGCACA  |                          |
| 1101         | <del>-</del>             |                          |                          | CGTCAGATCG<br>GCAGTCTAGC  |                          |
| 1151         |                          |                          |                          | ACACCGGGAC<br>TGTGGCCCTG  |                          |
| 1201         |                          |                          |                          | GGATTCCCCG<br>CCTAAGGGGC  |                          |
| 1251         |                          |                          |                          | CAAGAGGTCC<br>GTTCTCCAGG  |                          |
| 1301         |                          |                          |                          | CCGAGCCCGC<br>GGCTCGGGCG  |                          |
| 1351         |                          |                          |                          | GTGGGCGCCG<br>_CACCCGCGCC |                          |
| 1401         | ••••                     |                          |                          | CAACACCGCC<br>GTTGTGGCGG  |                          |
| 1451         |                          |                          |                          | ACGAGGAGGT<br>TGCTCCTCCA  |                          |
| 1501         |                          |                          |                          | ACCTACAAGG<br>TGGATGTTCC  |                          |
| 1551         |                          |                          |                          |                           | CTGATCCACT<br>GACTAGGTGA |
| 1601         | CCCAGAAGAG<br>GGGTCTTCTC |                          |                          |                           | CACCCAGGGC<br>GTGGGTCCCG |
| <b>16</b> 51 | TACTTCCCCG<br>ATGAAGGGGC |                          |                          |                           | TCAGGTTCCC<br>AGTCCAAGGG |
| 1701         | CCTGACCTTC<br>GGACTGGAAG |                          |                          |                           | CCCGAGAAGG<br>GGGCTCTTCC |
| 1751         | TGGAGGAGGC<br>ACCTCCTCCG | CAACGAGGGC<br>GTTGCTCCCG | GAGAACAACT<br>CTCTTGTTGA | GCGCCGCCA<br>CGCGGCGGGT   | CCCCATGTCC               |

Figure 27B

| 1801 |             |                |               |                          | GGAGGT TA<br>CCTCCAAGCT  |
|------|-------------|----------------|---------------|--------------------------|--------------------------|
| 1851 |             |                |               | GGAGCTGCAC               | CCCGAGTACT<br>GGGCTCATGA |
|      | GAGGTTCGAC  | CGGAAGGTGG     | TGCACCGGTC    | CCTCGACGTG               | GGGCTCATGA               |
| 1901 |             |                |               | CTGTGCCTTC<br>GACACGGAAG |                          |
|      | IGIICCIGAC  | GH111CGGGC     | CCGICIAGAC    | GACACGGAAAG              | AICANCOGIC               |
| 1951 |             |                |               | TCCTTGACCC               |                          |
|      |             |                | •             |                          |                          |
| 2001 |             |                |               | GGAAATTGCA               | TCGCATTGTC               |
|      |             |                |               |                          | -                        |
| 2051 |             |                |               | GGGTGGGGCA<br>CCCACCCCGT | CCTGTCGTTC               |
|      | 0000100100  | 0001101011     |               |                          | 0.00000000000            |
| 2101 |             |                |               | CGACCCCTAC               | CGGTGGGCTC<br>GCCACCCGAG |
| 2151 | m>mccccc nm | P000000000     | ma cmca a amo | TGTGGGCGTG               |                          |
| 2151 |             |                |               | ACACCCGCAC               |                          |
| 2201 | CCCNNNCNNC  | እጥኔ ጥል አ ር ርጥር | CCCCTCTTAT    | ലസ് നേനന്ന <b>െ</b>      | ATCTGTTTTG               |
| 2201 |             |                |               | CATCAAAACA               |                          |
| 2251 | ראפראפררפר  | רכררכרראיזיכ   | ACCACCAACT    | CGTTTGATGG               | ል ልርድ ልጥጥር ጥር            |
| 2232 |             |                |               | GCAAACTACC               |                          |
| 2301 | AGCTCATATT  | TGACAACGCG     | CATGCCCCCA    | TGGGCCGGGG               | TGCGTCAGAA               |
|      |             |                |               |                          | ACGCAGTCTT               |
| 2351 | TGTGATGGGC  | TCCAGCATTG     | ATGGTCGCCC    | CGTCCTGCCC               | GCAAACTCTA               |
|      | ACACTACCCG  | AGGTCGTAAC     | TACCAGCGGG    | GCAGGACGGG               | CGTTTGAGAT               |
| 2401 | CTACCTTGAC  | CTACGAGACC     | GTGTCTGGAA    | CGCCGTTGGA               | GACTGCAGCC               |
|      | GATGGAACTG  | GATGCTCTGG     | CACAGACCTT    | GCGGCAACCT               | CTGACGTCGG               |
| 2451 |             |                |               |                          | TTGTGACTGA               |
|      | AGGCGGCGGC  | GAAGTCGGCG     | ACGTCGGTGG    | CGGGCGCCCT               | AACACTGACT               |
| 2501 |             |                |               |                          | CGTTCATCCG               |
|      | GAAACGAAAG  | GACTCGGGCG     | AACGTTTGTC    | ACGTCGAAGG               | GCAAGTAGGC               |
| 2551 | CCCGCGATGA  |                |               |                          | TTCTTTGACC<br>AAGAAACTGG |
|      |             |                |               |                          |                          |
| 2601 |             |                |               |                          | GCCAGCAGGT<br>CGGTCGTCCA |
|      |             |                |               |                          |                          |
| 2651 |             |                |               |                          | AACATAAATA<br>TTGTATTTAT |
|      |             |                |               |                          |                          |
| 2701 |             |                |               |                          | TTGCTGTCTT<br>AACGACAGAA |
|      |             |                |               |                          |                          |

Figure 27C

| 2751 | TATTTAGGGG               |                          |                          |                          |                          |
|------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
|      |                          |                          |                          | GCCCTGGTCG               |                          |
| 2801 | GTTGAGGGTC<br>CAACTCCCAG |                          |                          | GTGGTAAAGG<br>CACCATTTCC |                          |
| 2851 |                          |                          |                          | TGGGGTGGAG               |                          |
| 2031 |                          |                          |                          | ACCCCACCTC               |                          |
| 2901 |                          |                          |                          | TAGATGATCC<br>ATCTACTAGG |                          |
| 2951 |                          |                          |                          | TTTCAGTAGC<br>AAAGTCATCG |                          |
| 3001 |                          |                          |                          | CAAAGCGGTT<br>GTTTCGCCAA |                          |
| 3051 |                          |                          |                          | TTGGACTGTA<br>AACCTGACAT | TTTTTAGGTT<br>AAAAATCCAA |
| 3101 |                          |                          |                          | ATTCATGTTG<br>TAAGTACAAC | TGCAGAACCA<br>ACGTCTTGGT |
| 3151 |                          | •                        |                          | ATTTGTCATG<br>TAAACAGTAC | TAGCTTAGAA<br>ATCGAATCTT |
| 3201 |                          |                          |                          | TTGTGACCTC<br>AACACTGGAG |                          |
| 3251 |                          |                          |                          | CCCACGGGCG<br>GGGTGCCCGC |                          |
| 3301 |                          |                          |                          |                          | CAGGATGAGA<br>GTCCTACTCT |
| 3351 |                          |                          |                          | CGGAGGGTGC<br>GCCTCCCACG | CAGACTGCGG<br>GTCTGACGCC |
| 3401 |                          |                          |                          |                          | CAGATTTGCA<br>GTCTAAACGT |
| 3451 |                          |                          |                          |                          | CTGCGGGGCG<br>GACGCCCCGC |
| 3501 |                          |                          |                          |                          | AAGAAAGCAG<br>TTCTTTCGTC |
| 3551 | GTTCCTGAGC<br>CAAGGACTCG | AGCTGCGACT<br>TCGACGCTGA | TACCGCAGCC<br>ATGGCGTCGG | GCTGGGCCCG<br>CCACCCGGGC | TAAATCACAC<br>ATTTAGTGTG |
| 3601 |                          |                          |                          |                          | GCCGTCATCC<br>CGGCAGTAGG |
| 3651 | CTGAGCAGGG<br>GACTCGTCCC | GGGCCACTTC<br>CCCGGTGAAG | GTTAAGCATG<br>CAATTCGTAC | TCCCTGACTC<br>AGGGACTGAG | GCATGTTTTC<br>CGTACAAAAG |

Figure 270

| 3701 | CCTGACCAAA<br>GGACTGGTTT | GCCAGAA<br>AGGCGGTCTT    | GGCGCTCGCC<br>CCGCGAGCGG | GCCCAGCGAT<br>CGGGTCGCTA | AGCAGT TT<br>TCGTCAAGAA  |
|------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| 3751 |                          |                          | AACGGTTTGA<br>TTGCCAAACT |                          | CGTAGGCATG<br>GCATCCGTAC |
| 3801 | CTTTTGAGCG<br>GAAAACTCGC |                          | CAGTTCCAGG<br>GTCAAGGTCC |                          |                          |
| 3851 |                          |                          | CCAGCATATC<br>GGTCGTATAG |                          |                          |
| 3901 |                          |                          | GTAGTCGGTG<br>CATCAGCCAC |                          |                          |
| 3951 |                          |                          | AGGGTCCTCG<br>TCCCAGGAGC |                          |                          |
| 4001 |                          |                          |                          |                          | GCTTGAGGCT<br>CGAACTCCGA |
| 4051 |                          |                          | GCTGCCGGTC<br>CGACGGCCAG |                          | GCGTCGGCCA<br>CGCAGCCGGT |
| 4101 |                          |                          |                          |                          | GGCGTGGCCC               |
| 4151 |                          |                          |                          |                          | GGCAGTGCAG<br>CCGTCACGTC |
| 4201 |                          |                          |                          |                          | TCCGGGGAGT<br>AGGCCCCTCA |
| 4251 |                          |                          |                          |                          | CACGAGCCAG<br>GTGCTCGGTC |
| 4301 |                          |                          |                          |                          | CATGCTTTTT<br>GTACGAAAAA |
| 4351 |                          |                          |                          | _                        | CGCTCGCTGA<br>GCGAGCCACT |
| 4401 |                          |                          |                          |                          | CCTGTCCTCG<br>GGACAGGAGC |
| 4451 | AGCGGTGTTC<br>TCGCCACAAG | CGCGGTCCTC               | CTCGTATAGA<br>GAGCATATCT | AACTCGGACC<br>TTGAGCCTGG | ACTCTGAGAC<br>TGAGACTCTG |
| 4501 |                          |                          |                          |                          | GAGGGGTAGC<br>CTCCCCATCG |
| 4551 |                          |                          |                          |                          | AAGACACATG<br>TTCTGTGTAC |
| 4601 | TCGCCCTCTT<br>AGCGGGAGAA | CGCCATCAAG<br>GCCGTAGTTC | GAAGGTGATT<br>CTTCCACTAA | GGTTTGTAGG<br>CCAAACATCC | TGTAGGCCAC<br>ACATCCGGTG |

Figure 27E

| 4651 |                          | CAAGGACTTC               |      |                          |
|------|--------------------------|--------------------------|------|--------------------------|
| 4701 |                          | CTCTTCCGCA<br>GAGAAGGCGT | <br> |                          |
| 4751 |                          | TCTGAAAAGC<br>AGACTTTTCG | <br> |                          |
| 4801 |                          | GAGGAGGATT<br>CTCCTCCTAA |      |                          |
| 4851 |                          | CGCATCCATC<br>GCGTAGGTAG |      |                          |
| 4901 |                          | CAAACGACCC<br>GTTTGCTGGG |      |                          |
| 4951 |                          | GTTTGGTTTT<br>CAAACCAAAA |      |                          |
| 5001 | TGTTTAGCTG<br>ACAAATCGAC | CACGTATTCG<br>GTGCATAAGC | <br> |                          |
| 5051 |                          | CGTCGGGCAC<br>GCAGCCCGTG | <br> | '                        |
| 5101 | GGTGACAAGG<br>CCACTGTTCC | TCAACGCTGG<br>AGTTGCGACC |      |                          |
| 5151 |                          | CGCCGGCGG                | <br> |                          |
| 5201 |                          | CGTCCGGGGG<br>GCAGGCCCCC |      |                          |
| 5251 |                          | TCGAAGTAGT<br>AGCTTCATCA | <br> | TCTAGCGCCT<br>AGATCGCGGA |
| 5301 |                          |                          |      | GAGTGGGGGA<br>CTCACCCCCT |
| 5351 |                          |                          |      | CGCAAATGTC<br>GCGTTTACAG |
| 5401 |                          |                          |      | GGGTAGCATC<br>CCCATCGTAG |
| 5451 |                          |                          |      | GTGCGAGGGA<br>CACGCTCCCT |
| 5501 |                          |                          |      | CTGCTCGGAA<br>GACGAGCCTT |
| 5551 |                          |                          |      | GTTGGACGCT<br>CAACCTGCGA |

Figure 27F

| 5601 | GGAAGACGTT<br>CCTTCTGCAA | CTYCGACCGC               | TCTGTGAGAC<br>AGACACTCTG | CTACCGCGTC<br>GATGGCGCAG | ACGCAC G                 |
|------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| 5651 | GAGGCGTAGG<br>CTCCGCATCC | AGTCGCGCAG<br>TCAGCGCGTC | CTTGTTGACC<br>GAACAACTGG | AGCTCGGCGG<br>TCGAGCCGCC | TGACCTGCAC<br>ACTGGACGTG |
| 5701 |                          |                          |                          | GATGATGTCA<br>CTACTACAGT |                          |
| 5751 |                          |                          |                          | GGACAAACTC<br>CCTGTTTGAG |                          |
| 5801 | TTCCAGTACT<br>AAGGTCATGA | CTTGGATCGG<br>GAACCTAGCC | AAACCCGTCG<br>TTTGGGCAGC | GCCTCCGAAC<br>CGGAGGCTTG | GGTAAGAGCC<br>CCATTCTCGG |
| 5851 | ATCGTACATC               | TTGACCAACT               | GCCGGACCAT               | GGCGCAGCAT<br>CCGCGTCGTA | GGGAAAAGAT               |
| 5901 | GCCCATCGCG               | CATACGGACG               | CGCCGGAAGG               | GGAGCGAGGT<br>CCTCGCTCCA | CACCCACTCG               |
| 5951 | CGTTTCCACA               | GGGACTGGTA               | CTGAAACTCC               |                          | ACTTCAGTCA               |
| 6001 | CAGCAGCGTA               | GGCGGGACGA               | GGGTCTCGTT               | AAAGTCCGTG<br>TTTCAGGCAC | GCGAAAAACC               |
| 6051 | TTGCGCCTAA               | ACCGTCCCGC               | TTCCACTGTA               | GCAACTTCTC               | TATCTTTCCC<br>ATAGAAAGGG |
| 6101 | CGCGCTCCGT               | ATTTCAACGC               | ACACTACGCC               | TTCCCAGGGC               | GCACCTCGGA<br>CGTGGAGCCT |
| 6151 | TGCCAACAAT               | TAATGGACCC               | GCCGCTCGTG               | CTAGAGCAGT               | AAGCCGTTGA<br>TTCGGCAACT |
| 6201 | ACAACACCGG               | GTGTTACATT               | TCAAGGTTCT               | TCGCGCCCTA               | GCCCTTGATG<br>CGGGAACTAC |
| 6251 | CTTC:CGTTAA              | AAAATTCAAG               | GAGCATCCAC               | TCGAGAAGTC               | GGGAGCTGAG<br>CCCTCGACTC |
| 6301 | GGGCACGAGA               | CTTTCCCGGG               | TCAGACGTTC               | TACTCCCAAC               | GAAGCGACGA<br>CTTCGCTGCT |
|      | TACTCGAGGT               | GTCCAGTGCC               | CGGTAATCGT               | AAACGTCCAC               | GTCGCGAAAG<br>CAGCGCTTTC |
|      | CAGGATTTGA               | CCGCTGGATA               | CCGGTAAAAA               | AGACCCCACT               | TGCAGTAGAA<br>ACGTCATCTT |
|      | CCATTCGCCC               | AGAACAAGGG               | TCGCCAGGGT               | AGGTTCCAAG               | GCGGCTAGGT               |
| 6501 | CTCGCGCGGC<br>GAGCGCGCCG | AGTCACTAGA<br>TCAGTGATCT | GGCTCATCTC               | CGCCGAACTT<br>GCGGCTTGAA | CATGACCAGC<br>GTACTGGTCG |

Figure 27G

| 6551  | ATGAAGGGCA<br>TACTTCCCGT | GCTCGACGAA | CCCAAAGGCC<br>GGGTTTCCGG | CCCATCCAAG<br>GGGTAGGTTC | TATAGE C<br>ATATCCAGAG   |
|-------|--------------------------|------------|--------------------------|--------------------------|--------------------------|
| 6601  |                          |            |                          | GCGAGGATGC<br>CGCTCCTACG |                          |
| 6651  | •                        |            | •                        | AGGAGTGGCT<br>TCCTCACCGA |                          |
| 6701  |                          |            |                          | CACTCGTGCT<br>GTGAGCACGA |                          |
| 6751  |                          |            |                          | GGGCTGTACA<br>CCCGACATGT |                          |
| 6801  |                          |            |                          | AGAGTGGGAA<br>TCTCACCCTT |                          |
| 6851  |                          |            |                          | ACTTCGGCTG<br>TGAAGCCGAC |                          |
| 6901  |                          |            |                          | GGATCGGACC<br>CCTAGCCTGG |                          |
| 6951  |                          |            |                          | GCGGTCGGAG<br>CGCCAGCCTC |                          |
| 7001  |                          |            |                          | TGGAGCTCCC<br>ACCTCGAGGG |                          |
| 7051  |                          |            |                          | GCATAGACGG<br>CGTATCTGCC |                          |
| 7101  |                          |            |                          | GGGGCTGGTT<br>CCCCGACCAA | -                        |
| 7151  |                          |            | •                        | GGCGCGACTA<br>CCGCGCTGAT |                          |
| 7201  |                          |            |                          | GGATGATGCA<br>CCTACTACGT | TCTAAAAGCG<br>AGATTTTCGC |
| 7251  |                          |            | •                        | GGGCTCCGGA<br>CCCGAGGCCT |                          |
| 7301  | GAGGGGGCAG<br>CTCCCCCGTC |            |                          |                          | CTGGTGCTGC<br>GACCACGACG |
| 7351  | GCGCGTAGGT<br>CGCGCATCCA |            |                          |                          | TCTCCTGAAT<br>AGAGGACTTA |
| 7401  | CTGGCGCCTC<br>GACCGCGGAG |            |                          |                          | AACCTGAAAG<br>TTGGACTTTC |
| 7,451 |                          |            |                          |                          | CTGGCGCAAA<br>GACCGCGTTT |

Figure 27H

| 7501 | ATCTCCTGCA<br>TAGAGGACGT | CTCTGA<br>GAGGACT        | GTTGTCTTGA<br>CAACAGAACT | TAGGCGATOR<br>ATCCGCTAGA | GCCGGT. TT               |
|------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| 7551 |                          | TCTTCCTCCT<br>AGAAGGAGGA |                          |                          |                          |
| 7601 |                          | GTCGTTGGAA<br>CAGCAACCTT |                          |                          |                          |
| 7651 | <del>-</del>             | CGTTCCAGAC<br>GCAAGGTCTG |                          |                          |                          |
| 7701 | ••••                     | ATGACCACCT<br>TACTGGTGGA |                          |                          |                          |
| 7751 |                          | GTTTCGCAGG<br>CAAAGCGTCC |                          |                          |                          |
| 7801 |                          | CCACGAAGAA<br>GGTGCTTCTT |                          |                          |                          |
| 7851 |                          | CCCAAGGCCT<br>GGGTTCCGGA |                          |                          |                          |
| 7901 |                          | GAAAAACTGG<br>CTTTTTGACC |                          |                          |                          |
| 7951 | TCCAGAAGAC<br>AGGTCTTCTG | GGATGAGCTC<br>CCTACTCGAG | GGCGACAGTG<br>CCGCTGTCAC | TCGCGCACCT<br>AGCGCGTGGA | CGCGCTCAAA<br>GCGCGAGTTT |
| 8001 |                          | GCCTCTTCTT<br>CGGAGAAGAA |                          |                          |                          |
| 8051 |                          | TTCTTCTGGC<br>AAGAAGACCG |                          |                          | ACGGCGGCGA<br>TGCCGCCGCT |
| 8101 |                          | CCGGGAGGCG<br>GGCCCTCCGC |                          |                          |                          |
| 8151 |                          | ATGGTCTCGG<br>TACCAGAGCC |                          |                          | CGGGGGCGCA<br>GCCCCCGCGT |
| 8201 |                          | GCCGCCCGTC<br>CGGCGGGCAG |                          |                          | CGGGGGGCTG<br>GCCCCCGAC  |
| 8251 | CCATGCGGCA<br>GGTACGCCGT | GGGATACGGC<br>CCCTATGCCG | GCTAACGATG<br>CGATTGCTAC | CATCTCAACA<br>GTAGAGTTGT | ATTGTTGTGT<br>TAACAACACA |
| 8301 | AGGTACTCCG<br>TCCATGAGGC | CCGCCGAGGG<br>GGCGGCTCCC | ACCTGAGCGA<br>TGGACTCGCT | GTCCGCATCG<br>CAGGCGTAGC | ACCGGATCGG<br>TGGCCTAGCC |
| 8351 | AAAACCTCTC<br>TTTTGGAGAG | GAGAAAGGCG<br>CTCTTTCCGC | TCTAACCAGT<br>AGATTGGTCA | CACAGTCGCA<br>GTGTCAGCGT | AGGTAGGCTG<br>TCCATCCGAC |
| 8401 | AGCACCGTGG<br>TCGTGGCACC | CGGCCGCCAG<br>GCCCGCCGTC | CGGCGCGCGC               | TCGGGGTTGT<br>AGCCCCAACA | TTCTGGCGGA<br>AAGACCGCCT |

Figure 27I

| 8451 | GGTGCTGCTG<br>CCACGACGAC | ATGTAAT<br>TACTACATTA    | TAAAGTAGGC<br>ATTTCATCCG | GGTCTTGAGA<br>CCAGAACTCT | CGGCGCTACC               |
|------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| 8501 |                          | CACCATGTCC<br>GTGGTACAGG |                          |                          |                          |
| 8551 | TCGGCCATGC<br>AGCCGGTACG | CCCAGGCTTC<br>GGGTCCGAAG |                          |                          |                          |
| 8601 |                          | AGCCTTTCTA<br>TCGGAAAGAT |                          |                          | TCCTCTTGTC<br>AGGAGAACAG |
| 8551 | CTGCATCTCT<br>GACGTAGAGA | TGCATCTATC<br>ACGTAGATAG |                          |                          |                          |
| 8701 |                          | TTCCTCCCAT<br>AAGGAGGGTA |                          |                          | TCATCGGCTG<br>AGTAGCCGAC |
| 8751 |                          | AGGTCGGCGA<br>TCCAGCCGCT |                          |                          | GCCTGCTGCA<br>CGGACGACGT |
| 8801 |                          |                          |                          |                          | GCGGTGGTAT<br>CGCCACCATA |
| 8851 |                          | TGATGGTGTA<br>ACTACCACAT |                          |                          | ACCAGTTAAC<br>TGGTCAATTG |
| 8901 |                          | CCCGGCTGCG<br>GGGCCGACGC |                          |                          | CGCGAGTAAG<br>GCGCTCATTC |
| 8951 |                          |                          |                          |                          | GTACTGGTAT<br>CATGACCATA |
| 9001 |                          |                          |                          |                          | AGCGTAGGGT<br>TCGCATCCCA |
| 9051 |                          |                          |                          |                          | TGATATCCGT<br>ACTATAGGCA |
| 9101 |                          | GGACATCCAG<br>CCTGTAGGTC |                          |                          |                          |
| 9151 |                          |                          |                          |                          | AAAAGTGCTC<br>TTTTCACGAG |
| 9201 |                          |                          |                          |                          | TTGACGCTCT<br>AACTGCGAGA |
| 9251 |                          |                          |                          |                          | GTGGTCTGGT<br>CACCAGACCA |
| 9301 |                          |                          |                          |                          | CGAGCCCCGT<br>GCTCGGGGCA |
| 9351 |                          |                          |                          |                          | TGTCGAACCC<br>ACAGCTTGGG |

Figure 27J

| 9401      |                          | GEAGTETGTT               |      |                          |
|-----------|--------------------------|--------------------------|------|--------------------------|
| 9451      |                          | CTGCTGCGCT<br>GACGACGCGA |      |                          |
| 9501      | TAAGCGGTTA<br>ATTCGCCAAT | GGCTGGAAAG<br>CCGACCTTTC |      |                          |
| 9551<br>· |                          | ATTTTCCAAG<br>TAAAAGGTTC | <br> |                          |
| 9601      |                          | CGGACTGCGG<br>GCCTGACGCC |      |                          |
| 9651      |                          | GCAAATTCCT<br>CGTTTAAGGA |      |                          |
| 9701      | TTTCCCAGAT<br>AAAGGGTCTA | GCATCCGGTG<br>CGTAGGCCAC |      |                          |
| 9751      |                          | AAGAGCAGCG<br>TTCTCGTCGC |      |                          |
| 9801      | TACCGCGTCA<br>ATGGCGCAGT | GGAGGGGGGA<br>CCTCCCCGCT |      |                          |
|           | ATTACGAACC<br>TAATGCTTGG |                          |      |                          |
| 9901      |                          | TGGCGCGGCT<br>ACCGCGCCGA |      |                          |
| 9951      |                          | AAGCGTGATA<br>TTCGCACTAT |      |                          |
| 10001     |                          | CCGCGAGGGA<br>GGCGCTCCCT |      |                          |
|           | TTCCACGCAG<br>AAGGTGCGTC |                          |      |                          |
| 10101     |                          | GACTTTGAGC<br>CTGAAACTCG |      | AGTCCCGCGC<br>TCAGGGCGCG |
| 10151     |                          |                          |      | GCAGACGGTG<br>CGTCTGCCAC |
| 10201     |                          |                          |      | TGCGTACGCT<br>ACGCATGCGA |
| 10251     |                          |                          |      | TGGGACTTTG<br>ACCCTGAAAC |
| 10301     |                          |                          |      | GGCGCAGCTG<br>CCGCGTCGAC |

Figure 27K

| 10351 |                          | T GCACAG<br>ACGTCGTGTC   |            |            |                          |
|-------|--------------------------|--------------------------|------------|------------|--------------------------|
| 10401 |                          | GTAGAGCCCG<br>CATCTCGGGC |            |            |                          |
| 10451 |                          | CATAGTGGTG<br>GTATCACCAC | •          |            |                          |
| 10501 |                          | TCAACTATTC<br>AGTTGATAAG |            |            |                          |
| 10551 |                          | CATACCCCTT<br>GTATGGGGAA |            |            |                          |
| 10601 |                          | CATGCGCATG<br>GTACGCGTAC |            |            |                          |
| 10651 |                          | ATCGCAACGA<br>TAGCGTTGCT |            |            |                          |
| 10701 |                          | CTCAGCGACC<br>GAGTCGCTGG |            |            |                          |
| 10751 |                          | GGGCAGCGGC<br>CCCGTCGCCG |            |            |                          |
| 10801 |                          | TGCGCTGGGC<br>ACGCGACCCG |            |            |                          |
| 10851 |                          | GGGCTGGCGG<br>CCCGACCGCC |            |            |                          |
| 10901 |                          | ATATGACGAG<br>TATACTGCTC |            |            |                          |
| 10951 |                          | TGATGTTTCT<br>ACTACAAAGA |            |            |                          |
| 11001 |                          | CGGCGCTGCA<br>GCCGCGACGT |            |            |                          |
|       |                          | GTCCAGTACC               | TGGCGTAGTA | CAGCGACTGA | CGCGCGTTAG               |
| 11101 |                          |                          |            |            | CGCAATTCTG<br>GCGTTAAGAC |
|       |                          | AGGGCCGCGC               | GCGTTTGGGG | TGCGTGCTCT | TCCACGACCG               |
| 11201 |                          |                          |            |            | GACGAGGCCG<br>CTGCTCCGGC |
| 11251 | GCCTGGTCTA<br>CGGACCAGAT |                          |            |            | CAACAGCGGC<br>GTTGTCGCCG |

Figure 27L

| 11301 | AACGTGCAGA               | CCTGGA                   | CCGGCTGGTG               | GGGGATGTGC               | GCGAGC T                 |
|-------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
|       | •                        |                          |                          | CCCCTACACG               |                          |
| 11351 | GGCGCAGCGT               | GAGCGCGCGC               | AGCAGCAGGG               | CAACCTGGGC               | TCCATGGTTG               |
|       | CCGCGTCGCA               | CTCGCGCGCG               | TCGTCGTCCC               | GTTGGACCCG               | AGGTACCAAC               |
| 11401 | CACTAAACGC               | CTTCCTGAGT               | ACACAGCCCG               | CCAACGTGCC               | GCGGGGACAG               |
|       | GTGATTTGCG               | GAAGGACTCA               | TGTGTCGGGC               | GGTTGCACGG               | CGCCCTGTC                |
| 11451 |                          |                          |                          | CGGCTAATGG<br>GCCGATTACC |                          |
| 11501 | ACCGCAAAGT               | GAGGTGTACC               | AGTCTGGGCC               | AGACTATTTT               | TTCCAGACCA               |
|       | TGGCGTTTCA               | CTCCACATGG               | TCAGACCCGG               | TCTGATAAAA               | AAGGTCTGGT               |
| 11551 | GTAGACAAGG               | CCTGCAGACC               | GTAAACCTGA               | GCCAGGCTTT               | CAAAAACTTG               |
|       | CATCTGTTCC               | GGACGTCTGG               | CATTTGGACT               | CGGTCCGAAA               | GTTTTTGAAC               |
| 11601 | CAGGGGCTGT               | GGGGGGTGCG               | GGCTCCCACA               | GGCGACCGCG               | CGACCGTGTC               |
|       | GTCCCCGACA               | CCCCCCACGC               | CCGAGGGTGT               | CCGCTGGCGC               | GCTGGCACAG               |
| 11651 | TAGCTTGCTG               | ACGCCCAACT               | CGCGCCTGTT               | GCTGCTGCTA               | ATAGCGCCCT               |
|       | ATCGAACGAC               | TGCGGGTTGA               | GCGCGGACAA               | CGACGACGAT               | TATCGCGGGA               |
| 11701 | TCACGGACAG               | TGGCAGCGTG               | TCCCGGGACA               | CATACCTAGG               | TCACTTGCTG               |
|       | AGTGCCTGTC               | ACCGTCGCAC               | AGGGCCCTGT               | GTATGGATCC               | AGTGAACGAC               |
| 11751 | ACACTGTACC               | GCGAGGCCAT               | AGGTCAGGCG               | CATGTGGACG               | AGCATACTTT               |
|       | TGTGACATGG               | CGCTCCGGTA               | TCCAGTCCGC               | GTACACCTGC               | TCGTATGAAA               |
| 11801 | CCAGGAGATT               | ACAAGTGTCA               | GCCGCGCGCT               | GGGGCAGGAG               | GACACGGGCA               |
|       | GGTCCTCTAA               | TGTTCACAGT               | CGGCGCGCGA               | CCCCGTCCTC               | CTGTGCCCGT               |
| 11851 | GCCTGGAGGC               | AACCCTAAAC               | TACCTGCTGA               | CCAACCGGCG               | GCAGAAGATC               |
|       | CGGACCTCCG               | TTGGGATTTG               | ATGGACGACT               | GGTTGGCCGC               | CGTCTTCTAG               |
| 11901 | CCCTCGTTGC               | ACAGTTTAAA               | CAGCGAGGAG               | GAGCGCATTT               | TGCGCTACGT               |
|       | GGGAGCAACG               | TGTCAAATTT               | GTCGCTCCTC               | CTCGCGTAAA               | ACGCGATGCA               |
| 11951 | GCAGCAGAGC               | GTGAGCCTTA               | ACCTGATGCG               | CGACGGGGTA               | ACGCCCAGCG               |
|       | CGTCGTCTCG               | CACTCGGAAT               | TGGACTACGC               | GCTGCCCCAT               | TGCGGGTCGC               |
| 12001 | TGGCGCTGGA<br>ACCGCGACCT | CATGACCGCG               | CGCAACATGG<br>GCGTTGTACC | AACCGGGCAT<br>TTGGCCCGTA | GTATGCCTCA<br>CATACGGAGT |
| 12051 | AACCGGCCGT<br>TTGGCCGGCA | TTATCAACCG               | CCTAATGGAC<br>GGATTACCTG | TACTTGCATC<br>ATGAACGTAG | CCCCCCCCC<br>CCCCCCCCC   |
| 12101 | CGTGAACCCC               | GAGTATITCA<br>CTCATAAAGT | CCAATGCCAT<br>GGTTACGGTA | CTTGAACCCG<br>GAACTTGGGC | CACTGGCTAC<br>GTGACCGATG |
| 12151 | CGCCCCTGG                | TTTCTACACC               | GGGGGATTCG<br>CCCCTAAGC  | AGGTGCCCGA<br>TCCACGGGCT | GGGTAACGAT<br>CCCATTGCTA |
| 12201 | GGATTCCTCT               | GGGACGACAT               | AGACGACAGC               | GTGTTTTCCC               | CGCAACCGCA               |
|       | CCTAAGGAGA               | CCCTGCTGTA               | TCTGCTGTCG               | CACAAAAGGG               | GCGTTGGCGT               |

Figure 27 M

| 12251 |                        |                   |                   | GGCAGAGGCG |                     |
|-------|------------------------|-------------------|-------------------|------------|---------------------|
|       | CTGGGACGAT             | CAACGTTG          | TOGOGOTOGT        | CCGTCTCCGC | CGCGAC              |
|       |                        |                   |                   |            |                     |
| 12301 | ACCABACCTT             | CCCCA CCCCA       | a coa common      | CCGATCTAGG | COMPONE             |
| 12301 |                        |                   |                   |            |                     |
|       | TCCTTTCGAA             | GGCGTCCGGT        | TUGTUGAACA        | GGCTAGATCC | GUGAUGUUGG          |
|       |                        |                   |                   |            |                     |
| 12351 | CCGCGGTCAG             | ATGCTAGTAG        | CCCATTTCCA        | AGCTTGATAG | GGTCTCTTAC          |
|       | GGCGCCAGTC             | TACGATCATC        | GGGTAAAGGT        | TCGAACTATC | CCAGAGAATG          |
|       |                        |                   |                   |            |                     |
| 12401 | CAGCACTCGC             | ACCACCCCC         | CCCCCTCCT         | GGGCGAGGAG | СУСТУССТУУ          |
|       |                        |                   |                   | CCCGCTCCTC |                     |
|       | G1CG1GAGCG             | 196199666         | GCGCGGACGA        | CCCGCTCCTC | CICAIGGAIT          |
|       |                        |                   |                   |            |                     |
| 12451 |                        |                   |                   | AAAACCTGCC |                     |
|       | TGTTGAGCGA             | CGACGTCGGC        | GTCGCGCTTT        | TTTTGGACGG | AGGCCGTAAA          |
|       |                        |                   |                   | •          |                     |
| 12501 | CCCAACAACG             | GGATAGAGAG        | CCTAGTGGAC        | AAGATGAGTA | GATGGAAGAC          |
|       |                        |                   |                   | TTCTACTCAT |                     |
|       |                        |                   | 0011010010        |            | C1110C11C1G         |
| 10551 | CE10000010             | 010010100         |                   |            |                     |
| 12551 |                        |                   |                   | ccccccccc  |                     |
|       | CATGCGCGTC             | CTCGTGTCCC        | TGCACGGTCC        | GGGCGCGGGC | GGGTGGGCAG          |
|       |                        |                   |                   |            |                     |
| 12601 | GTCAAAGGCA             | CGACCGTCAG        | CGGGGTCTGG        | TGTGGGAGGA | CGATGACTCG          |
|       | CAGTTTCCGT             | GCTGGCAGTC        | GCCCCAGACC        | ACACCCTCCT | GCTACTGAGC          |
|       | •                      |                   |                   |            |                     |
| 12651 | CCACACCACA             | CCACCCMCCM        | COMMODOCO         | GGGAGTGGCA | 30000mmm00          |
| 12031 |                        |                   |                   |            |                     |
|       | CGTCTGCTGT             | CGTCGCAGGA        | CCTAAACCCT        | CCCTCACCGT | TGGGCAAACG          |
|       |                        |                   |                   |            |                     |
| 12701 | GCACCTTCGC             | CCCAGGCTGG        | GGAGAATGTT        | TTAAAAAAAA | AAAAAGCATG          |
|       | CGTGGAAGCG             | GGGTCCGACC        | CCTCTTACAA        | TTTTTTTAA  | TTTTTCGTAC          |
|       |                        |                   |                   |            |                     |
| 12751 | ስጥርር እ <b>አ አ አጥ</b> አ | אאאארתראר         | CANGGCCATG        | GCACCGAGCG | <b>ጥጥር/ርጥጥጥጥር</b> ጥ |
| 12,31 |                        |                   |                   | CGTGGCTCGC |                     |
|       | IACGIIIIAI             | TITITIGAGIG       | GITCCGGTAC.       | CG1GGC1CGC | MACLAMAMGA          |
|       |                        |                   | •                 |            |                     |
| 12801 |                        |                   |                   | ATGTATGAGG |                     |
|       | ACATAAGGGG             | AATCATACGC        | CGCGCGCCGC        | TACATACTCC | TTCCAGGAGG          |
|       |                        |                   |                   |            |                     |
| 12851 | TCCCTCCTAC             | GAGAGTGTGG        | TGAGCGCGGC        | GCCAGTGGCG | GCGGCGCTGG          |
|       |                        |                   |                   | CGGTCACCGC |                     |
|       | AGGGAGGA1 G            | C1C1CACACC        | ACTOGGGGGG        | CGGICACCGC | COCCOCOACC          |
|       |                        |                   |                   |            |                     |
| 12901 |                        |                   |                   | CGTTTGTGCC |                     |
|       | CAAGAGGGAA             | GCTACGAGGG        | GACCTGGGCG        | GCAAACACGG | AGGCGCCATG          |
|       |                        |                   |                   |            |                     |
| 12951 | CTGCGGCCTA             | CCGGGGGGAG        | AAACAGCATC        | CGTTACTCTG | AGTTGGCACC          |
|       | CACCCCCCAT             | CCCCCCCCC         | THYCHCCHAC        | GCAATGAGAC | TCAACCGTGG          |
|       |                        | 000000000         |                   |            | 1011111111111       |
| 12001 | 0071770000             | > 00 > 00 D D D D | mam. comean       |            | maxxaaaxaa          |
| 13001 |                        |                   |                   |            |                     |
|       | GGATAAGCTG             | TGGTGGGCAC        | ACATGGACCA        | CCTGTTGTTC | AGTIGCCIAC          |
|       |                        |                   |                   |            |                     |
| 13051 | TGGCATCCCT             | GAACTACCAG        | AACGACCACA        | GCAACTTTCT | GACCACGGTC          |
|       |                        |                   |                   | CGTTGAAAGA |                     |
| •     |                        |                   |                   |            |                     |
| 12101 | איייעעעעעעע            | ATTCA CTIACA C    | רירות מייים מייים | CCNACCACAC | AGACCATCAA          |
| T3101 |                        |                   |                   |            |                     |
|       | TAAGTTTTGT             | TACTGATGIC        | GGGGGGGGGG        | CGTTCGTGTG | TUTGGTAGTT          |
|       |                        |                   |                   |            |                     |
| 13151 | TCTTGACGAC             |                   |                   |            |                     |
|       | AGAACTGCTG             | GCCAGCGTGA        | CCCCGCCGCT        | GGACTTTTGG | TAGGACGTAT          |

Figure 27N

| 13201 | CCAACATGCC               | A GTGAAC                 | GAGTTCATGT               | TTACCAATAA ~             | GTTTARFORG               |
|-------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
|       | GGTTGTACGG               | T. CACTTG                | CTCAAGTACA               | AATGGTTATT               | CAAATT                   |
| 13251 | CGGGTGATGG               | TGTCGCGCTT               | GCCTACTAAG               | GACAATCAGG               | TGGAGCTGAA               |
|       | GCCCACTACC               | ACAGCGCGAA               | CGGATGATTC               | CTGTTAGTCC               | ACCTCGACTT               |
| 13301 |                          |                          |                          | GGGCAACTAC<br>CCCGTTGATG |                          |
| 13351 | TGACCATAGA               | CCTTATGAAC               | AACGCGATCG               | TGGAGCACTA               | CTTGAAAGTG               |
|       | ACTGGTATCT               | GGAATACTTG               | TTGCGCTAGC               | ACCTCGTGAT               | GAACTTTCAC               |
| 13401 | GGCAGACAGA               | ACGGGGTTCT               | GGAAAGCGAC               | ATCGGGGTAA               | AGTTTGACAC               |
|       | CCGTCTGTCT               | TGCCCCAAGA               | CCTTTCGCTG               | TAGCCCCAȚT               | TCAAACTGTG               |
| 13451 | CCGCAACTTC               | AGACTGGGGT               | TTGACCCCGT               | CACTGGTCTT               | GTCATGCCTG               |
|       | GCCGTTGAAG               | TCTGACCCCA               | AACTGGGGCA               | GTGACCAGAA               | CAGTACGGAC               |
| 13501 | GGGTATATAC               | AAACGAAGCC               | TTCCATCCAG               | ACATCATTTT               | GCTGCCAGGA               |
|       | CCCATATATG               | TTTGCTTCGG               | AAGGTAGGTC               | TGTAGTAAAA               | CGACGGTCCT               |
| 13551 | TGCGGGGTGG               | ACTTCACCCA               | CAGCCGCCTG               | AGCAACTTGT               | TGGGCATCCG               |
|       | ACGCCCACC                | TGAAGTGGGT               | GTCGGCGGAC               | TCGTTGAACA               | ACCCGTAGGC               |
| 13601 | CAAGCGGCAA               | CCCTTCCAGG               | AGGGCTTTAG               | GATCACCTAC               | GATGATCTGG               |
|       | GTTCGCCGTT               | GGGAAGGTCC               | TCCCGAAATC               | CTAGTGGATG               | CTACTAGACC               |
| 13651 | AGGGTGGTAA               | CATTCCCGCA               | CTGTTGGATG               | TGGACGCCTA               | CCAGGCGAGC               |
|       | TCCCACCATT               | GTAAGGGCGT               | GACAACCTAC               | ACCTGCGGAT               | GGTCCGCTCG               |
| 13701 | TTGAAAGATG<br>AACTTTCTAC | ACACCGAACA<br>TGTGGCTTGT | CCCCCCCA                 | GGCGCAGGCG<br>CDCCTDCDCC | GCAGCAACAG<br>CGTCGTTGTC |
| 13751 | CAGTGGCAGC<br>GTCACCGTCG | GGCGCGGAAG<br>CCGCGCCTTC | AGAACTCCAA<br>TCTTGAGGTT | CGCGCCACCC               | GCGGCAATGC<br>CGCCGTTACG |
| 13801 |                          |                          |                          | TTCGCGGCGA<br>AAGCGCCGCT |                          |
| 13851 |                          |                          |                          | GAAGCAGCGG<br>CTTCGTCGCC |                          |
| 13901 | CGCCCCGCT                | GCGCAACCCG               | AGGTCGAGAA               | GCCTCAGAAG               | AAACCGGTGA               |
|       | GCGGGGGGGA               | CGCGTTGGGC               | TCCAGCTCTT               | CGGAGTCTTC               | TTTGGCCACT               |
| 13951 | TCAAACCCCT               | GACAGAGGAC               | AGCAAGAAAC               | GCAGTTACAA               | CCTAATAAGC               |
|       | AGTTTGGGGA               | CTGTCTCCTG               | TCGTTCTTTG               | CGTCAATGTT               | GGATTATTCG               |
| 14001 | AATGACAGCA<br>TTACTGTCGT | CCTTCACCCA<br>GGAAGTGGGT | CATGGCGTCG               | TGGTACCTTG<br>ACCATGGAAC | CATACAACTA<br>GTATGTTGAT |
| 14051 | CGGCGACCCT<br>GCCGCTGGGA | CAGACCGGAA<br>GTCTGGCCTT | TCCGCTCATG               | GACCCTGCTT<br>CTGGGACGAA | TGCACTCCTG<br>ACGTGAGGAC |
| 14101 | ACGTAACCTG<br>TGCATTGGAC | CGGCTCGGAG               | CAGGTCTACT               | GGTCGTTGCC               | AGACATGATG<br>TCTGTACTAC |

Tigure 270

WO 02/22080

| 14151 | CAAGACCCCG<br>GTTCTGGGGC  | TYTTCCG<br>ACTGGAAGGC    | CTCCACGCGC<br>GAGGTGCGCG | CAGATCAGCA<br>GTCTAGTCGT | ACTTTC T<br>TGAAAGGCCA   |
|-------|---------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| 14201 |                           | GAGCTGTTGC<br>CTCGACAACG |                          |                          |                          |
| 14251 | AGGCCGTCTA<br>TCCGGCAGAT  | CTCCCAACTC<br>GAGGGTTGAG | ATCCGCCAGT<br>TAGGCGGTCA | TTACCTCTCT<br>AATGGAGAGA | GACCCACGTG<br>CTGGGTGCAC |
| 14301 |                           | TTCCCGAGAA<br>AAGGGCTCTT |                          |                          |                          |
| 14351 |                           | GTCAGTGAAA<br>CAGTCACTTT |                          |                          |                          |
| 14401 |                           | CAACAGCATC<br>GTTGTCGTAG |                          |                          |                          |
| 14451 |                           | GCACCTGCCC<br>CGTGGACGGG |                          |                          |                          |
| 14501 |                           | CTATCGAGCC<br>GATAGCTCGG |                          |                          |                          |
| 14551 |                           | CAATAACACA<br>GTTATTGTGT |                          |                          |                          |
| 14601 |                           | CCAAGAAGCG<br>GGTTCTTCGC |                          |                          |                          |
| 14651 |                           | GCGCCCTGGG<br>CGCGGGACCC |                          |                          |                          |
| 14701 |                           | TGACGCCATC<br>ACTGCGGTAG |                          |                          |                          |
| 14751 |                           | CGCCACCAGT<br>GCGGTGGTCA |                          |                          |                          |
| 14801 | GGTGCGCGGA<br>'CCACGCGCCT | GCCCGGCGCT<br>CGGGCCGCGA |                          |                          |                          |
| 14851 |                           | CCACCGCCGC<br>GCTGGCGGCG |                          |                          |                          |
| 14901 |                           | TTAACCGCGC<br>AATTGGCGCG |                          |                          |                          |
| 14951 |                           | AGGCTGGCCG<br>TCCGACCGGC |                          |                          | CCCAGGTCCA<br>GGGTCCAGGT |
| 15001 | GGCGACGAGC<br>CCGCTGCTCG  |                          |                          |                          | TATGACTCAG<br>ATACTGAGTC |
| 15051 |                           |                          |                          |                          | GCGGCCTGCG<br>CGCCGGACGC |

Figure 27P

| 15101 | CCTCCCCCTG       | CCCCCCC     | CCCCGCGCAA  | CTAGATTGCA               | AGAAAA     |
|-------|------------------|-------------|-------------|--------------------------|------------|
|       | GCACGGGCAC       | esta ececce | GGGGCGCGTT  | GATCTAACGT               | TCTTTTT    |
| 15151 |                  |             |             | CGGCGGCGGC               |            |
|       |                  |             |             | . ecceccece              |            |
| 15201 |                  |             |             | ATGCTCCAGG               |            |
|       | CGATACAGGT       | TCGCGTTTTA  | GTTTCTTCTC  | TACGAGGTCC               | AGTAGCGCGG |
| 15251 |                  |             |             | GCAGGATTAC               |            |
|       |                  |             |             | CGTCCTAATG               |            |
| 15301 |                  |             |             | ATGATGATGA               |            |
|       |                  |             | •           | TACTACTACT               |            |
| 15351 |                  |             |             | CCCAGGCGAC               |            |
|       |                  |             |             | GGCTCCGCTG               |            |
| 15401 | GAAAGGTCGA       | CGCGTAAAAC  | GTGTTTTGCG  | ACCCGGCACC               | ACCGTAGTCT |
|       |                  |             |             | TGGGCCGTGG               |            |
| 15451 |                  | _           |             | ACAAGCGCGT               |            |
|       |                  |             |             | TGTTCGCGCA               |            |
| 15501 |                  |             |             | GCCAACGAGC               |            |
|       |                  |             |             | CGGTTGCTCG               |            |
| 15551 |                  |             |             | GCTGGCGTTG               |            |
|       |                  |             |             | CGACCGCAAC               |            |
| 15601 |                  |             |             | TAACACTGCA<br>ATTGTGACGT |            |
|       |                  |             | -           | GGCCTAAAGC               |            |
| 15651 | CCCGCGCTTG       | CTCCCTCCCA  | ACAMAMGC GC | CCGGATTTCG               | CCCTCAGACC |
| 15701 |                  |             |             | ACCCAAGCGC               |            |
| 13701 |                  |             |             | TEGETTCECE               |            |
|       | no i dianece o x |             | 200.2111001 |                          |            |
| 15751 | AAGATGTCTT       | GGAAAAAATG  | ACCGTGGAAC  | CTGGGCTGGA               | GCCCGAGGTC |
|       | TTCTACAGAA       | CCTTTTTTAC  | TCCCACCTTC  | GACCCGACCT               | CGGGCTCCAG |
| 15801 |                  |             |             |                          | TGCAGACCGT |
|       |                  |             |             |                          | ACGTCTGGCA |
| 15851 | GGACGTTCAG       | ATACCCACTA  | CCAGTAGCAC  | CAGTATTGCC               | ACCGCCACAG |
|       |                  |             |             |                          | TEGEGGTETC |
| 15901 | AGGGCATGGA       | GACACAAACG  | TCCCCCCTTC  | CCTCAGCGGT               | GGCGGATGCC |
|       |                  |             | _           |                          | CCGCCTACGG |
| 15951 | GCGGTGCAGG       | CGGTCGCTGC  | GGCCGCGTCC  | AAGACCTCTA               | CGGAGGTGCA |
|       |                  |             |             |                          | GCCTCCACGT |
| 16001 | AACGGACCCG       | TGGATGTTTC  | GCGTTTCAGC  | CCCCCGGCGC               | CCGCGCCGTT |
|       | TTGCCTGGGC       | ACCTACAAAG  | CGCAAAGTCG  | GGGGCCGCG                | GGCGCGGCAA |

Figure 270

| 16051       |              | CCCCCCC                  |             |            |                          |
|-------------|--------------|--------------------------|-------------|------------|--------------------------|
|             | GCTCCTTCAT   | GCCGCGGCGG               | TCGCGCGATG  | ACGGGCTTAT | ACGGGATGTA               |
| 16101       | CCTTCCATTG   | CGCCTACCCC               | CGGCTATCGT  | GGCTACACCT | ACCGCCCCAG               |
|             | GGAAGGTAAC   | GCGGATGGGG               | GCCGATAGCA  | CCGATGTGGA | TGGCGGGGTC               |
| 16151       | AAGACGAGCA   | ACTACCCGAC               | GCCGAACCAC  | CACTGGAACC | CGCCGCCGCC               |
|             | TTCTGCTCGT   | TGATGGGCTG               | CGGCTTGGTG  | GTGACCTTGG | GCGGCGGCGG               |
| 16201 .     | GTCGCCGTCG   | CCAGCCCGTG               | CTGGCCCCGA  | TTTCCGTGCG | CAGGGTGGCT               |
|             | CAGCGGCAGC   | GGTCGGGCAC               | GACCGGGGCT  | AAAGGCACGC | GTCCCACCGA               |
| 16251       | CGCGAAGGAG   | GCAGGACCCT               | GGTGCTGCCA  | ACAGCGCGCT | ACCACCCCAG               |
|             |              | CGTCCTGGGA               |             |            |                          |
| 16301       | CATCGTTTAA   | AAGCCGGTCT               | TTGTGGTTCT  | TGCAGATATG | GCCCTCACCT               |
|             |              | TTCGGCCAGA               |             |            |                          |
| 16351       | GCCGCCTCCG   | TTTCCCGGTG               | CCGGGATTCC  | GAGGAAGAAT | GCACCGTAGG               |
|             |              | AAAGGGCCAC               |             |            |                          |
| 16401       | ACCCCCATCC   | CCGGCCACGG               |             | GGCATGCGTC | GTGCGCACCA               |
| 10401       |              | GGCCGGTGCC               |             |            |                          |
| 16451       | ccecce       | CGCGCGTCGC               | አሮሮርማሮርሮስጥ  | CCCCCCCCT  | איזירריזיפרררר           |
| 10431       |              | GCGCGCAGCG               |             |            |                          |
| 16501       | mccmm x mmcc | N CIDCN TICCC            | CCCCCAPTC   |            | CGGAATTGCA               |
| 10201       |              | TGACTAGCGG               |             |            |                          |
| 16551       | maaamaaaam   | maa. aaaaa.              | C) C) C) C) | mmaaaaa.   | CDDCC MCDC               |
| 16551       |              | ACGTCCGCGT               |             |            | GTTGCATGTG<br>CAACGTACAC |
|             |              |                          |             |            |                          |
| 16601       |              | TTATTTTTCA               |             |            | GGTCCTGTAA               |
| •           |              |                          |             | •          |                          |
| 16651       |              | GAATGGAAGA<br>CTTACCTTCT | -           |            | CCCCGCGACA               |
|             | GATAAAACAT   | CTIACCTICT               | GIAGIIGAAA  | CGCAGAGACC | 66666666                 |
| 16701       |              |                          |             |            | ACCAGCAATA               |
|             | GCCGAGCGCG   | GGCAAGTACC               | CTTTGACCGT  | TCTATAGCCG | TGGTCGTTAT               |
| 16751       |              |                          |             |            | CATTAAAAAT               |
|             | ACTCGCCACC   | GCGGAAGTCG               | ACCCCGAGCG  | ACACCTCGCC | GTAATTTTTA               |
| 16801       |              |                          |             |            | ACAGCAGCAC               |
|             | AAGCCAAGGT   | GGCAATTCTT               | GATACCGTCG  | TTCCGGACCT | TGTCGTCGTG               |
| 16851       |              |                          |             |            | CAACAAAAGG               |
|             | TCCGGTCTAC   | GACTCCCTAT               | TCAACTITCT  | CGTTTTAAAG | GTTGTTTTCC               |
| 16901       |              |                          |             |            | CCTGGCCAAC               |
|             | ACCATCTACC   | GGACCGGAGA               | CCGTAATCGC  | CCCACCACCT | GGACCGGTTG               |
| 16951       | CAGGCAGTGC   | AAAATAAGAT               | TAACAGTAAG  | CTTGATCCCC | GCCCTCCCGT               |
| <del></del> |              |                          |             |            | CGGGAGGGCA               |

Figure 27R

| 17001 | AGAGGAGCCT<br>TCTCCTCGGA | cereccec<br>careccec                    | TGGAGACAGT<br>ACCTCTGTCA | GTCTCCAGAG<br>CAGAGGTCTC | CCCCCACCEC               |
|-------|--------------------------|---|--------------------------|--------------------------|--------------------------|
| 17051 |                          | GCGCCCCGAC<br>CGCGGGGCTG                |                          |                          |                          |
| 17101 |                          | CGTACGAGGA<br>GCATGCTCCT                |                          |                          |                          |
| 17151 |                          | CCCATGGCTA<br>GGGTACCGAT                |                          |                          |                          |
| 17201 |                          | GCCTCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC |                          |                          |                          |
| 17251 |                          | CCGTTGTTGT<br>GGCAACAACA                |                          |                          |                          |
| 17301 |                          | GGTCCGCGAT<br>CCAGGCGCTA                |                          |                          |                          |
| 17351 |                          | GAACAGCATC<br>CTTGTCGTAG                |                          |                          |                          |
| 17401 |                          | TCTGATAGCT<br>AGACTATCGA                |                          |                          |                          |
| 17451 |                          | CCAGAGGAGC<br>GGTCTCCTCG                | ,                        |                          |                          |
| 17501 |                          | CTTCGATGAT<br>GAAGCTACTA                |                          |                          |                          |
| 17551 |                          | TCGGAGTACC<br>AGCCTCATGG                |                          |                          |                          |
| 17601 |                          | GTACTTCAGC<br>CATGAAGTCG                |                          |                          |                          |
| 17651 |                          | ACGACGTGAC<br>TGCTGCACTG                |                          |                          |                          |
| 17701 |                          | GTGGACCGTG<br>CACCTGGCAC                |                          |                          |                          |
| 17751 |                          | TGTGGGTGAT<br>ACACCCACTA                |                          |                          | TTCCACGTAC<br>AAGGTGCATG |
| 17801 |                          |   |                          |                          | AGCCCTACTC<br>TCGGGATGAG |
| 17851 | TGGCACTGCC<br>ACCGTGACGG | TACAACGCCC<br>ATGTTGCGGG                | TGGCTCCCAA<br>ACCGAGGGTT | GGGTGCCCCA<br>CCCACGGGGT | AATCCTTGCG<br>TTAGGAACGC |
| 17901 |                          |   |                          |                          | AGAAGAGGAC<br>TCTTCTCCTG |

Figure 275

| 17951 | •                        | _                                     | GCTGAGCAGC<br>CGACTCGTCG     |                          |
|-------|--------------------------|---------------------------------------|------------------------------|--------------------------|
| 18001 |                          |                                       | <br>AAATATTACA<br>TTTATAATGT |                          |
| 18051 | TTCAAATAGG<br>AAGTTTATCC |                                       | <br>AATATGCCGA<br>TTATACGGCT |                          |
| 18101 |                          | · · · · · · · · · · · · · · · · · · · | <br>TGGTACGAAA<br>ACCATGCTTT |                          |
| 18151 |                          |                                       | TACCCCAATG<br>ATGGGGTTAC     |                          |
| 18201 |                          |                                       | <br>ATGGAGGGCA<br>TACCTCCCGT |                          |
| 18251 |                          |                                       | CAAGTGGAAA<br>GTTCACCTTT     |                          |
| 18301 |                          |                                       | TGATAACTTG<br>ACTATTGAAC     |                          |
| 18351 |                          |                                       | <br>AAACCCCAGA<br>TTTGGGGTCT |                          |
| 18401 |                          | •                                     | <br>TCACGAGAAC<br>AGTGCTCTTG |                          |
| 18451 |                          |                                       | <br>TGCTTTTAGG<br>ACGAAAATCC |                          |
| 18501 |                          |                                       | ATATGGGTGT<br>TATACCCACA     |                          |
| 18551 |                          |                                       | TTGCAAGACA<br>AACGTTCTGT     | GAAACACAGA<br>CTTTGTGTCT |
| 18601 |                          |                                       | TGGTGATAGA<br>ACCACTATCT     |                          |
| 18651 |                          |                                       | ATGATCCAGA<br>TACTAGGTCT     |                          |
| 18701 |                          |                                       |                              | GCTTTCCACT<br>CGAAAGGTGA |
| 18751 |                          |                                       |                              | CCTAAAACAG<br>GGATTTTGTC |
| 18801 |                          |                                       |                              | AGATAAAAAT<br>TCTATTTTTA |
| 18851 |                          |                                       |                              | TAAATGCCAA<br>ATTTACGGTT |

Figure 27T

| 18901  | CCTGTGGAGA               |                          |                          | AGCGCTGTAT<br>TCGCGACATA |                          |
|--------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| 10051  | AGCTAAAGTA               | •                        |                          |                          |                          |
| 18951  |                          |                          |                          | AAAGACTATT               |                          |
| 19001  |                          |                          |                          | CCCGGGCTAG<br>GGGCCCGATC |                          |
| 19051  |                          |                          |                          | CTATATGGAC               |                          |
|        |                          |                          |                          | GATATACCTG               |                          |
| 19101  |                          |                          |                          | GCTACCGCTC<br>CGATGGCGAG |                          |
| 19151  |                          |                          |                          | CAGGTGCCTC<br>GTCCACGGAG |                          |
|        |                          |                          |                          | CTCATACACC               |                          |
| 19201  | ACGGTAATTT               | TTGGAGGAAG               | AGGACGGCCC               | GAGTATGTGG               | ATGCTCACCT               |
| 19251  |                          |                          |                          | AGAGCTCCCT<br>TCTCGAGGGA |                          |
| 19301  | CTAAGGGTTG               | ACGGAGCCAG               | CATTAAGTTT               | GATAGCATTT               | GCCTTTACGC               |
|        |                          |                          |                          | CTATCGTAAA               |                          |
| 19351  | CACCTTCTTC<br>GTGGAAGAAG | CCCATGGCCC<br>GGGTACCGGG | ACAACACCGC<br>TGTTGTGGCG | CTCCACGCTT<br>GAGGTGCGAA | GAGGCCATGC<br>CTCCGGTACG |
| 19401  | TTAGAAACGA<br>AATCTTTGCT | CACCAACGAC<br>GTGGTTGCTG | CAGTCCTTTA<br>GTCAGGAAAT | ACGACTATCT<br>TGCTGATAGA | CTCCGCCGCC<br>GAGGCGGCGG |
| 19451  | AACATGCTCT               | ACCCTATACC               | CGCCAACGCT               | ACCAACGTGC               | CCATATCCAT<br>GGTATAGGTA |
| 10501  |                          |                          |                          |                          | ACGCGCCTTA               |
| 19501  | CCCTCCCC                 | TTGACCCGCC               | GAAAGGCGCC               | GACCCGGAAG               | TGCGCGGAAT               |
| 19551. | AGACTAAGGA<br>TCTGATTCCT | AACCCCATCA<br>TTGGGGTAGT | CTGGGCTCGG<br>GACCCGAGCC | GCTACGACCC<br>CGATGCTGGG | TTATTACACC<br>AATAATGTGG |
| 19601  |                          |                          |                          |                          | TCAACCACAC               |
|        |                          |                          |                          |                          | AGTTGGTGTG               |
| 19651  | CTTTAAGAAG<br>GAAATTCTTC | GTGGCCATTA<br>CACCGGTAAT | GGAAACTGAG               | AAGACAGTCG               | TGGCCTGGCA<br>ACCGGACCGT |
| 19701  | ATGACCGCCT<br>TACTGGCGGA | GCTTACCCCC<br>CGAATGGGGG | AACGAGTTTG<br>TTGCTCAAAC | AAATTAAGCG<br>TTTAATTCGC | CTCAGTTGAC<br>GAGTCAACTG |
| 19751  | GGGGAGGGTT<br>CCCCTCCCAA | ACAACGTTGC<br>TGTTGCAACG | CCAGTGTAAC<br>GGTCACATTG | ATGACCAAAG<br>TÄCTGGTTTC | ACTGGTTCCT<br>TGACCAAGGA |
| 19801  | GGTACAAATG<br>CCATGTTTAC | CTAGCTAACT<br>GATCGATTGA | ATAACATIGG<br>TATIGTAACC | CTACCAGGGC<br>GATGGTCCCG | TTCTATATCC<br>AAGATATAGG |

Figure 274

| 19851   | CAGAGAGCTA<br>GTCTCTCGAT | CL GACCGC<br>GTTCCTGGCG  | ATGTACTCCT<br>TACATGAGGA | TCTTTAGAAA<br>AGAAATCTTT | CTTCCA(CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC |
|---------|--------------------------|--------------------------|--------------------------|--------------------------|---|
| 19901   | ATGAGCCGTC<br>TACTCGGCAG | AGGTGGTGGA<br>TCCACCACCT | TGATACTAAA<br>ACTATGATTT | TACAAGGACT<br>ATGTTCCTGA | ACCAACAGGT<br>TGGTTGTCCA                |
| 19951   |                          | CACCAACACA<br>GTGGTTGTGT |                          |                          |   |
| - 20001 | CCACCATGCG<br>GGTGGTACGC | CGAAGGACAG<br>GCTTCCTGTC | GCCTACCCTG<br>CGGATGGGAC | CTAACTTCCC<br>GATTGAAGGG | CTATCCGCTT<br>GATAGGCGAA                |
| 20051   | ATAGGCAAGA<br>TATCCGTTCT | CCGCAGTTGA<br>GGCGTCAACT | CAGCATTACC<br>GTCGTAATGG | CAGAAAAAGT<br>GTCTTTTCA  | TTCTTTGCGA<br>AAGAAACGCT                |
| 20101   | TCGCACCCTT<br>AGCGTGGGAA | TGGCGCATCC<br>ACCGCGTAGG | CATTCTCCAG<br>GTAAGAGGTC | TAACTTTATG<br>ATTGAAATAC | TCCATGGGCG<br>AGGTACCCGC                |
| 20151   | CACTCACAGA<br>GTGAGTGTCT | CCTGGGCCAA<br>GGACCCGGTT | AACCTTCTCT<br>TTGGAAGAGA | ACGCCAACTC<br>TGCGGTTGAG | CGCCCACGCG<br>GCGGGTGCGC                |
| 20201   | CTAGACATGA<br>GATCTGTACT | CTTTTGAGGT<br>GAAAACTCCA | GGATCCCATG<br>CCTAGGGTAC | GACGAGCCCA<br>CTGCTCGGGT | CCCTTCTTTA<br>GGGAAGAAAT                |
| 20251   |                          | GAAGTCTTTG<br>CTTCAGAAAC |                          |                          |   |
| 20301   |                          | AACCGTGTAC<br>TTGGCACATG |                          |                          |   |
| 20351   |                          | GAAGCAAGCA<br>CTTCGTTCGT |                          |                          |   |
| 20401   |                          | AACTGAAAGC<br>TTGACTTTCG |                          |                          |   |
| 20451   |                          | ACCTATGACA<br>TGGATACTGT |                          |                          |   |
| 20501   |                          | CGCCATAGTC<br>GCGGTATCAG |                          |                          |   |
| 20551   |                          | CCTTTGCCTG<br>GGAAACGGAC |                          |                          |   |
| 20601   |                          | GGCTTTTCTG<br>CCGAAAAGAC |                          |                          |   |
| 20651   |                          | ACTCCTGCGC<br>TGAGGACGCG |                          |                          |   |
| 20701   | TGTATAACGC<br>ACATATTGCG | TGGAAAAGTC<br>ACCTTTTCAG |                          |                          |   |
| 20751   | CGCCTGTGGA<br>GCGGACACCT | CTATTCTGCT<br>GATAAGACGA |                          |                          |   |

Figure 27 V.

| 20801  | CCCAAACTCC<br>GGGTTTGAGG | GATCAC<br>GTACCTAGTG | AACCCCACCA<br>TTGGGGTGGT | TGAACCTTAT<br>ACTTGGAATA | TACCGG A                 |
|--------|--------------------------|----------------------|--------------------------|--------------------------|--------------------------|
| 20851  |                          |                      | TCCCCAGGTA<br>AGGGGTCCAT |                          |                          |
| 20901  |                          |                      | TCCTGGAGCG<br>AGGACCTCGC |                          |                          |
| 20,951 |                          |                      | AGCGCCACTT<br>TCGCGGTGAA |                          |                          |
| 21001  |                          |                      | AGACACTTTC<br>TCTGTGAAAG |                          |                          |
| 21051  | TTTGTACACT               | CTCGGGTGAT           | TATTTACCCC               | CACCCTTGCC               | GTCTGCGCCG               |
|        | AAACATGTGA               | GAGCCCACTA           | ATAAATGGGG               | GTGGGAACGG               | CAGACGCGGC               |
| 21101  | TTTAAAAATC               | AAAGGGGTTC           | TGCCGCGCAT               | CGCTATGCGC               | CACTGGCAGG               |
|        | AAATTTTTAG               | TTTCCCCAAG           | ACGGCGCGTA               | GCGATACGCG               | GTGACCGTCC               |
| 21151  | GACACGTTGC               | GATACTGGTG           | TTTAGTGCTC               | CACTTAAACT               | CAGGCACAAC               |
|        | CTGTGCAACG               | CTATGACCAC           | AAATCACGAG               | GTGAATTTGA               | GTCCGTGTTG               |
| 21201  | CATCCGCGGC               | AGCTCGGTGA           | AGTTTTCACT               | CCACAGGCTG               | CGCACCATCA               |
|        | GTAGGCGCCG               | TCGAGCCACT           | TCAAAAGTGA               | GGTGTCCGAC               | GCGTGGTAGT               |
| 21251  |                          |                      |                          |                          | GCAGTTGGGG<br>CGTCAACCCC |
| 21301  | CCTCCGCCCT               | GCGCGCGCGA           | GTTGCGATAC               | ACAGGGTTGC               | AGCACTGGAA               |
|        | GGAGGCGGGA               | CGCGCGCGCT           | CAACGCTATG               | TGTCCCAACG               | TCGTGACCTT               |
| 21351  | CACTATCAGC               | GCCGGGTGGT           | GCACGCTGGC               | CAGCACGCTC               | TTGTCGGAGA               |
|        | GTGATAGTCG               | CGGCCCACCA           | CGTGCGACCG               | GTCGTGCGAG               | AACAGCCTCT               |
| 21401  | TCAGATCCGC               | GTCCAGGTCC           | TCCGCGTTGC               | TCAGGGCGAA               | CGGAGTCAAC               |
|        | AGTCTAGGCG               | CAGGTCCAGG           | AGGCGCAACG               | AGTCCCGCTT               | GCCTCAGTTG               |
| 21451  |                          |                      |                          |                          | TTGAGTTGCA<br>AACTCAACGT |
| 21501  | CTCGCACCGT               | AGTGGCATCA           | AAAGGTGACC               | GTGCCCGGTC               | TGGGCGTTAG               |
|        | GAGCGTGGCA               | TCACCGTAGT           | TTTCCACTGG               | CACGGGCCAG               | ACCCGCAATC               |
| 21551  | GATACAGCGC               | CTGCATAAAA           | GCCTTGATCT               | GCTTAAAAGC               | CACCTGAGCC               |
|        | CTATGTCGCG               | GACGTATTTT           | CGGAACTAGA               | CGAATTTTCG               | GTGGACTCGG               |
| 21601  | TTTGCGCCTT               | CAGAGAAGAA           | CATGCCGCAA               | GACTTGCCGG               | AAAACTGATT               |
|        | AAACGCGGAA               | GTCTCTTCTT           | GTACGGCGTT               | CTGAACGGCC               | TTTTGACTAA               |
| 21651  | GGCCGGACAG               | GCCGCGTCGT           | GCACGCAGCA               | CCTTGCGTCG               | GTGTTGGAGA               |
|        | CCGGCCTGTC               | CGGCGCAGCA           | CGTGCGTCGT               | GGAACGCAGC               | CACAACCTCT               |
| 21701  | TCTGCACCAC               | ATTTCGGCCC           | CACCGGTTCT               | TCACGATCTT               | GGCCTTGCTA               |
|        | AGACGTGGTG               | TAAAGCCGGG           | GTGGCCAAGA               | AGTGCTAGAA               | CCGGAACGAT               |

Figure 27 W

| 21751 | GACTGCTCCT<br>CTGACGAGGA |      | TCGCTCGTCA<br>AGCGAGCAGT |                          |
|-------|--------------------------|------|--------------------------|--------------------------|
| 21801 |                          | -    | TCCGTGTAGA<br>AGGCACATCT |                          |
| 21851 |                          | <br> | ACAACGCGCA<br>TGTTGCGCGT |                          |
| 21901 |                          |      | GACTGCAGGT<br>CTGACGTCCA |                          |
| 21951 |                          |      | GTTGCTGGTG<br>CAACGACCAC |                          |
| 22001 |                          | <br> | TCTTGCATAC<br>AGAACGTATG |                          |
| 22051 | GCTTCCACTT<br>CGAAGGTGAA | <br> | TTCGCCTTTA<br>AAGCGGAAAT |                          |
| 22101 |                          |      | AGCCTCCATG<br>TCGGAGGTAC |                          |
| 22151 |                          | <br> | TCATCACCGT<br>AGTAGTGGCA |                          |
| 22201 |                          | <br> | TGCGTCCGCA<br>ACGCAGGCGT |                          |
| 22251 |                          |      | TGTGCGCTTA<br>ACACGCGAAT |                          |
| 22301 |                          |      | AACCCACCAT<br>TTGGGTGGTA |                          |
| 22351 |                          |      | ATTACCTCTG<br>TAATGGAGAC |                          |
| 22401 |                          |      | TTTCTTCTTG<br>AAAGAAGAAC |                          |
| 22451 |                          | <br> | GGCTGGGTGT<br>CCGACCCACA |                          |
| 22501 |                          |      |                          | TACGCCGCCT<br>ATGCGGCGGA |
| 22551 |                          |      |                          | GGGGACGGGG<br>CCCCTGCCCC |
| 22601 |                          | <br> |                          | GCGTCCGCGC<br>CGCAGGCGCG |
| 22651 |                          |      |                          | TTTCCTTCTC<br>AAAGGAAGAG |

Figure 27X

| 22701 | CTATAGGCAG | AGATCA       | TGGAGTCAGT | CGAGAAGAAG | GACAGC     |
|-------|------------|--------------|------------|------------|------------|
|       | GATATCCGTC | TTTTTCTAGT   | ACCTCAGTCA | GCTCTTCTTC | CTGTCGGATT |
| 22751 | CCGCCCCCTC |              |            |            |            |
|       | GGCGGGGGAG | ACTCAAGCGG   | TGGTGGCGGA | GGTGGCTACG | GCGGTTGCGC |
| 22801 |            | TCCCCGTCGA   |            |            |            |
|       | GGATGGTGGA | AGGGGCAGCT   | CCGTGGGGGC | GAACTCCTCC | TCCTTCACTA |
| 22851 | TATCGAGCAG |              |            |            |            |
|       | ATAGCTCGTC | CTGGGTCCAA   | AACATTCGCT | TCTGCTGCTC | CTGGCGAGTC |
| 22901 |            | GGATAAAAAG   |            |            |            |
|       | ATGGTTGTCT | CCTATTTTTC   | GTTCTGGTCC | TGTTGCGTCT | CCGTTTGCTC |
| 22951 |            | GGCGGGGGA    |            |            |            |
|       | CTTGTTCAGC | CCCCCCCCT    | GCTTTCCGTA | CCGCTGATGG | ATCTACACCC |
| 23001 |            | CTGTTGAAGC   |            |            |            |
|       | TCTGCTGCAC | GACAACTTCG   | TAGACGTCGC | GGTCACGCGG | TAATAGACGC |
| 23051 | ACGCGTTGCA | AGAGCGCAGC   | GATGTGCCCC | TCGCCATAGC | GGATGTCAGC |
|       | TGCGCAACGT | TCTCGCGTCG   | CTACACGGGG | AGCGGTATCG | CCTACAGTCG |
| 23101 |            |              |            |            | CCAAACGCCA |
|       | GAACGGATGC | TTGCGGTGGA   | TAAGAGTGGC | GCGCATGGGG | GGTTTGCGGT |
| 23151 |            |              |            |            | TACCCCGTAT |
|       | TCTTTTGCCG | TGTACGCTCG   | GGTTGGGCGC | GGAGTTGAAG | ATGGGGCATA |
| 23201 |            |              |            |            | CCAAAACTGC |
|       | AACGGCACGG | TCTCCACGAA   | CGGTGGATAG | TGTAGAAAAA | GGTTTTGACG |
| 23251 | AAGATACCCC | TATCCTGCCG   | TGCCAACCGC | AGCCGAGCGG | ACAAGCAGCT |
|       | TTCTATGGGG | ATAGGACGGC   | ACGGTTGGCG | TCGGCTCGCC | TGTTCGTCGA |
| 23301 |            |              |            |            | CTCAACGAAG |
|       | CCGGAACGCC | GTCCCGCGAC   | AGTATGGACT | ATAGCGGAGC | GAGTTGCTTC |
| 23351 | TGCCAAAAAT | CTTTGAGGGT   | CTTGGACGCG | ACGAGAAGCG | CGCGGCAAAC |
|       | ACGGTTTTTA | . GAAACTCCCA | GAACCTGCGC | TGCTCTTCGC | GCGCCGTTTG |
| 23401 |            |              |            |            | GAGTGTTGGT |
|       | CGAGACGTTG | TCCTTTTGTC   | GCTTTTACTT | TCAGTGAGAC | CTCACAACCA |
| 23451 | GGAACTCGAG | GGTGACAACG   | CGCGCCTAGC | CGTACTAAAA | CGCAGCATCG |
|       | CCTTGAGCTC | CCACTGTTGC   | GCGCGGATCG | GCATGATTTT | GCGTCGTAGC |
| 23501 | AGGTCACCCA | CTTTGCCTAC   | CCGGCACTTA | ACCTACCCC  | CAAGGTCATG |
|       | TCCAGTGGGT | GAAACGGATG   | GGCCGTGAAT | TGGATGGGGG | GTTCCAGTAC |
| 23551 | AGCACAGTCA | TGAGTGAGCT   | GATCGTGCGC | CGTGCGCAGC | CCCTGGAGAG |
|       | TCGTGTCAGT | ACTCACTCGA   | CTAGCACGCG | GCACGCGTCG | GGGACCTCTC |
| 23601 | GGATGCAAAT | TTGCAAGAAC   | AAACAGAGGA | GGGCCTACCC | GCAGTTGGCG |
|       | CCTACGTTTA | AACGTTCTTC   | TTTGTCTCCT | CCCGGATGGG | CGTCAACCGC |

Figure 27 Y

| 23651 | ACGAGCAGCT<br>TGCTCGTCGA  | A TOGOGOGACO             | CTTCAAACGC               | GCGAGCCTGC               | C CGACTT G               |
|-------|---------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| 23701 | GAGCGACGCA                | AACTAATGAT               | GGCCGCAGTG               | CTCGTTACCG               | TGGAGCTTGA               |
|       | CTCGCTGCG1                | TTGATTACTA               | CCGGCGTCAC               | GAGCAATGGC               | ACCTCGAACT               |
| 23751 | GTGCATGCAG<br>CACGTACGTC  | CGCTTCTTTG<br>CCCAAGAAAC | CTGACCCGGA<br>GACTGGGCCT | GATGCAGCGC CTACGTCGCG    | AAGCTAGAGG<br>TTCGATCTCC |
| 23801 | AAACATTGCA                | CTACACCTTT               | · CGACAGGGCT             | ACGTACGCCA               | GGCCTGCAAG               |
| 23851 |                           |                          |                          |                          | GAATTTTGCA               |
|       | TAGAGGTTGC                | ACCTCGAGAC               | GTTGGACCAG               | AGGATGGAAC               | CTTAAAACGT               |
| 23901 | CGAAAACCGC<br>GCTTTTGGCG  | CTTGGGCAAA<br>GAACCCGTTT | ACGTGCTTCA<br>TGCACGAAGT | TTCCACGCTC<br>AAGGTGCGAG | AAGGGCGAGG<br>TTCCCGCTCC |
| 23951 | CGCGCCGCGA<br>GCGCGCGCGCT | CTACGTCCGC<br>GATGCAGGCG | GACTGCGTTT<br>CTGACGCAAA | ACTTATTTCT               | ATGCTACACC<br>TACGATGTGG |
| 24001 | TGGCAGACGG                | CCATGGGCGT               | TTGGCAGCAG               | TGCTTGGAGG               | AGTGCAACCT               |
| 24053 |                           |                          |                          |                          | TCACGTTGGA               |
| 24021 | CAAGGAGCTG<br>GTTCCTCGAC  | GTCTTTGACG               | TAAAGCAAAA<br>ATTTCGTTTT | CTTGAAGGAC<br>GAACTTCCTG | CTATGGACGG<br>GATACCTGCC |
| 24101 | CCTTCAACGA<br>GGAAGTTGCT  | GCGCTCCGTG<br>CGCGAGGCAC | GCCGCGCACC<br>CGGCGCGTGG | TGGCGGACAT<br>ACCGCCTGTA | CATTTTCCCC<br>GTAAAAGGGG |
| 24151 | GAACGCCTGC                | TTAAAACCCT               | GCAACAGGGT               | CTGCCAGACT               | TCACCAGTCA               |
| 24201 |                           |                          | GGAACTTTAT               |                          |                          |
|       | TTCGTACAAC                | GTCTTGAAAT               | CCTTGAAATA               | GGATCTCGCG               | AGTCCTTAGA               |
| 24251 | TGCCCGCCAC<br>ACGGGCGGTG  | CTGCTGTGCA<br>GACGACACGT | CTTCCTAGCG<br>GAAGGATCGC | ACTTTGTGCC<br>TGAAACACGG | CATTAAGTAC<br>GTAATTCATG |
| 24301 | CGCGAATGCC<br>GCGCTTACGG  | CTCCGCCGCT<br>GAGGCGCGA  | TTGGGGCCAC<br>AACCCCGGTG | TGCTACCTTC<br>ACGATGGAAG | TGCAGCTAGC<br>ACGTCGATCG |
| 24351 |                           | GCCTACCACT               | CTGACATAAT<br>GACTGTATTA | GGAAGACGTG               | AGCGGTGACG               |
| 24401 | GTCTACTGGA                | GTGTCACTGT               | CGCTGCAACC               | TATGCACCCC               | GCACCGCTCC               |
|       | CAGATGACCT                | CACAGTGACA               | GCGACGTTGG               | ATACGTGGGG               | CGTGGCGAGG               |
| 24451 | CTGGTTTGCA<br>GACCAAACGT  | ATTCGCAGCT<br>TAAGCGTCGA | GCTTAACGAA<br>CGAATTGCTT | ACTCAAATTA<br>TCAGTTTAAT | TCGGTACCTT<br>AGCCATGGAA |
| 24501 | TGAGCTGCAG<br>ACTCGACGTC  | GGTCCCTCGC<br>CCAGGGAGCG | CTGACGAAAA<br>GACTGCTTTT | GTCCGCGGCT<br>CAGGCGCCGA | CCGGGGTTGA<br>GGCCCCAACT |
| 24551 | AACTCACTCC<br>TTGAGTGAGG  | GGGGCTGTGG               | ACGTCGGCTT               | ACCTTCGCAA               | ATTTGTACCT               |

Figure 27Z

| 24601 | GAGGACTACC<br>CTCCTGATGG | ACCCACGA<br>TGCGGGTGCT | CATTAGGTTC<br>CTAATCCAAG | TACGAAGACC<br>ATGCTTCTGG | AATCCCGGCCCCC            |
|-------|--------------------------|------------------------|--------------------------|--------------------------|--------------------------|
| 24651 |                          |                        | CCTGCGTCAT<br>GGACGCAGTA |                          |                          |
| 24701 | GCCAATTGCA               | AGCCATCAAC             | AAAGCCCGCC<br>TTTCGGGCGG | AAGAGTTTCT               | GCTACGAAAG               |
| 24751 | GGACGGGGG                | TTTACTTGGA             | CCCCCAGTCC               | GGCGAGGAGC               | TCAACCCAAT               |
| 24801 | CCCCCCCCCC               | CCGCAGCCCT             | ATCAGCAGCA<br>TAGTCGTCGT | ecceceecc                | CTTGCTTCCC               |
| 24851 | AGGATGGCAC               | CCAAAAAGAA             | GCTGCAGCTG<br>CGACGTCGAC | CCGCCGCCAC               | CCACGGACGA               |
| 24901 | GGAGGAATAC               | TGGGACAGTC             | AGGCAGAGGA<br>TCCGTCTCCT | GGTTTTGGAC               | GAGGAGGAGG               |
| 24951 | AGGACATGAT               | GGAAGACTGG             | GAGAGCCTAG<br>CTCTCGGATC | ACGAGGAAGC               | TTCCGAGGTC               |
| 25001 | GAAGAGGTGT               | CAGACGAAAC             | ACCGTCACCC<br>TGGCAGTGGG | TCGGTCGCAT               | TCCCCTCGCC               |
| 25051 | GGCGCCCCAG               | AAATCGGCAA             | CCGGTTCCAG<br>GGCCAAGGTC | CATGGCTACA               | ACCTCCGCTC               |
| 25101 | CTCAGGCGCC               | GCCGGCACTG             | CCCGTTCGCC<br>GGGCAAGCGG | GACCCAACCG               | TAGATGGGAC               |
| 25151 | ACCACTGGAA               | CCAGGGCCGG             | TAAGTCCAAG<br>ATTCAGGTTC | CAGCCGCCGC               | CGTTAGCCCA               |
| 25201 | AGAGCAACAA               | CAGCGCCAAG             | GCTACCGCTC               | ATGGCGCGGG               | CACAAGAACG               |
| 25251 | CCATAGTTGC               | TTGCTTGCAA             | CGATGGCGAG<br>GACTGTGGGG | GCAACATCTC               | CTTCGCCCGC               |
| 25301 | CGCTTTCTTC               | TCTACCATCA             |                          | TTCCCCCGTA               | ACATCCTGCA               |
| 25351 | TTACTACCGT               | CATCTCTACA             | GCCCATACTG               | CACCGGCGGC               | TGTAGGACGT               |
| 25401 | ACAGCAGCGG               | CCACACAGAA             | GCAAAGGCGA               | CCGGATAGCA               | TCGCCGTCGT<br>AGACTCTGAC |
| 25451 | AAAGCCCAAG               | AAATCCACAG             | CGGCGGCAGC               | AGCAGGAGGA               | TCTGAGACTG<br>GGAGCGCTGC |
|       | TTTCGGGTTC               | TTTAGGTGTC             | GCCGCCGTCG               | TCGTCCTCCT               | CCTCGCGACG<br>AAACAGGATT |
|       | CAGACCGCGG               | GTTGCTTGGG             | CATAGCTGGG               | CGCTCGAATC               | TTTGTCCTAA               |

Figure 27. AA

| 25551 | TTTCCCACTC<br>AAAGGGTGAG | TGCTAT<br>ACATACGATA     | ATTTCAACAG<br>TAAAGTTGTC | AGCAGGGGCC<br>TCGTCCCCGG | AAGAACA A<br>TTCTTGTTCT  |
|-------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| 25601 | GCTGAAAATA<br>CGACTTTTAT | AAAAACAGGT<br>TTTTTGTCCA |                          |                          |                          |
| 25651 |                          | CGAAGATCAG<br>GCTTCTAGTC |                          |                          |                          |
| 25701 |                          | AATACTGCGC<br>TTATGACGCG |                          |                          |                          |
| 25751 |                          | TAAGCGCGAA<br>ATTCGCGCTT |                          |                          |                          |
| 25801 |                          | TGTTGTCAGC<br>ACAAÇAGTCG |                          |                          |                          |
| 25851 |                          | GTTACCAGCC<br>CAATGGTCGG |                          |                          |                          |
| 25901 |                          | ACCCGAATAA<br>TGGGCTTATT |                          |                          |                          |
| 25951 |                          | CGGAATACGC<br>GCCTTATGCG |                          |                          |                          |
| 26001 |                          | CCACCACACC<br>GGTGGTGTGG |                          |                          |                          |
| 26051 |                          | GTGTACCAGG<br>CACATGGTCC |                          |                          |                          |
| 26101 |                          | CCAGGCCGAA<br>GGTCCGGCTT |                          |                          |                          |
| 26151 |                          | TTCGTCACAG<br>AAGCAGTGTC |                          |                          | •                        |
| 26201 |                          | AGAGGGCGAG<br>TCTCCCGCTC |                          |                          |                          |
| 26251 |                          | TCTCCGTCCG<br>AGAGGCAGGC |                          |                          |                          |
| 26301 | CGCTCTTCAT<br>GCGAGAAGTA | TCACGCCTCG<br>AGTGCGGAGC |                          |                          | _                        |
| 26351 | CTCTGAGCCG<br>GAGACTCGGC |                          |                          |                          | ATTGAGGAGT<br>TAACTCCTCA |
| 26401 | TTGTGCCATC<br>AACACGGTAG |                          |                          |                          | CGGCCACTAT<br>GCCGGTGATA |
| 26451 | CCGGATCAAT<br>GGCCTAGTTA | TTATTCCTAA<br>AATAAGGATT |                          |                          |                          |

Figure 27 AB

| 26501 | CTACGACTGA               | A TAAGTG                 | GAGAGGCAGA               | GCAACTGCGC               | CTGAAA C                 |
|-------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
|       | GATGCTGACT               | TACAATTCAC               | CTCTCCGTCT               | CGTTGACGCG               | GACTTTGTGG               |
| 26551 | TGGTCCACTG<br>ACCAGGTGAC |                          | AAGTGCTTTG<br>TTCACGAAAC |                          |                          |
| 26601 |                          |                          | GGATCATATC<br>CCTAGTATAG |                          |                          |
| 26651 |                          |                          | AGCTTGCCCG<br>TCGAACGGGC |                          |                          |
| 26701 |                          |                          | GAGCGGGACA<br>CTCGCCCTGT |                          |                          |
| 26751 | GTGATTTGCA               | ACTGTCCTAA               | CCCTGGATTA               | CATCAAGATC               | TTTGTTGCCA               |
|       | CACTAAACGT               | TGACAGGATT               | GGGACCTAAT               | GTAGTTCTAG               | AAACAACGGT               |
| 26801 | TCTCTGTGCT               | GAGTATAATA               | AATACAGAAA               | ATATATATA                | CTGGGGCTCC               |
|       | AGAGACACGA               | CTCATATTAT               | TTATGTCTTT               | TATATTTAA                | GACCCCGAGG               |
| 26851 | TATCGCCATC               | CTGTAAACGC               | CACCGTCTTC               | ACCCGCCCAA               | GCAAACCAAG               |
|       | ATAGCGGTAG               | GACATTTGCG               | GTGGCAGAAG               | TGGGCGGGTT               | CGTTTGGTTC               |
| 26901 | GCGAACCTTA               | CCTGGTACTT               | TTAACATCTC               | TCCCTCTGTG               | ATTTACAACA               |
|       | CGCTTGGAAT               | GGACCATGAA               | AATTGTAGAG               | AGGGAGACAC               | TAAATGTTGT               |
| 26951 | GTTTCAACCC               | AGACGGAGTG               | AGTCTACGAG               | AGAACCTCTC               | CGAGCTCAGC               |
|       | CAAAGTTGGG               | TCTGCCTCAC               | TCAGATGCTC               | TCTTGGAGAG               | GCTCGAGTCG               |
| 27001 | TACTCCATCA               | GAAAAAACAC               | CACCCTCCTT               | ACCTGCCGGG               | AACGTACGAG               |
|       | ATGAGGTAGT               | CTTTTTTGTG               | GTGGGAGGAA               | TGGACGGCCC               | TTGCATGCTC               |
| 27051 | TGCGTCACCG               | GCCGCTGCAC               | CACACCTACC               | GCCTGACCGT               | AAACCAGACT               |
|       | ACGCAGTGGC               | CGGCGACGTG               | GTGTGGATGG               | CGGACTGGCA               | TTTGGTCTGA               |
| 27101 | TTTTCCGGAC               | AGACCTCAAT               | AACTCTGTTT               | ACCAGAACAG               | GAGGTGAGCT               |
|       | AAAAGGCCTG               | TCTGGAGTTA               | TTGAGACAAA               | TGGTCTTGTC               | CTCCACTCGA               |
| 27151 | TAGAAAACCC               | TTAGGGTATT               | AGGCCAAAGG               | CGCAGCTACT               | GTGGGGTTTA               |
|       | ATCTTTTGGG               | AATCCCATAA               | TCCGGTTTCC               | GCGTCGATGA               | CACCCCAAAT               |
| 27201 | TGAACAATTC               | AAGCAACTCT               | ACGGGCTATT               | CTAATTCAGG               | TTTCTCTAGA               |
|       | ACTTGTTAAG               | TTCGTTGAGA               | TGCCCGATAA               | GATTAAGTCC               | AAAGAGATCT               |
| 27251 | ATCGGGGTTG               | GGGTTATTCT               | CTGTCTTGTG               | ATTCTCTTTA               | TTCTTATACT               |
|       | TAGCCCCAAC               | CCCAATAAGA               | GACAGAACAC               | TAAGAGAAAT               | AAGAATATGA               |
| 27301 | AACGCTTCTC               | TGCCTAAGGC               | TCGCCGCCTG               | CTGTGTGCAC               | ATTTGCATTT               |
|       | TTGCGAAGAG               | ACGGATTCCG               | AGCGGCGGAC               | GACACACGTG               | TAAACGTAAA               |
| 27351 | ATTGTCAGCT<br>TAACAGTCGA | TTTTAAACGC<br>AAAATTTGCG | TGGGGTCGCC               | ACCCAAGATG<br>TGGGTTCTAC | ATTAGGTACA<br>TAATCCATGT |
| 27401 | TAATCCTAGG<br>ATTAGGATCC | TTTACTCACC               | CTTGCGTCAG<br>GAACGCAGTC | CCCACGGTAC               | CACCCAAAAG<br>GTGGGTTTTC |

Ligure 27AC

| 27451 | GTGGATTTTA          | A GCCAGC           | CTGTAATGTT   | ACATTCGCAG            | CTGAAG                                  |
|-------|---------------------|--------------------|--------------|-----------------------|---|
|       | CACCTAAAAT          | TCCTCGGTCG         | GACATTACAA   | TGTAAGCGTC            | GACTTCGATT                              |
|       |                     |                    |              | ·                     |   |
| 27501 | TGAGTGCACC          | ACTOTTATAA         | AATGCACCAC   | AGAACATGAA            | AAGCTGCTTA                              |
|       | ACTCACGTGG          | TGAGAATATT         | TTACGTGGTG   | TCTTGTACTT            | TTCGACGAAT                              |
|       |                     |                    |              |                       |   |
| 27551 | TTCGCCACAA          | AAACAAAATT         | GGCAAGTATG   | CTGTTTATGC            | TATTTGGCAG                              |
|       | AAGCGGTGTT          | TTTGTTTTAA         | CCGTTCATAC   | GACAAATACG            | ATAAACCGTC                              |
|       |                     |                    |              |                       | *************************************** |
| 27601 | CCAGGTGACA          | CTACAGAGTA         | TAATGTTACA   | GTTTTCCAGG            | GTAAAACTCA                              |
|       |                     |                    | ATTACAATGT   |                       |   |
|       |                     |                    |              |                       |   |
| 27651 | TAAAACTTTT          | ATCTATACTT         | TTCCATTTTA   | <b>ምርኔ ኔ ኔ ጥርጥር</b> ር | CACATTACCA                              |
|       | ATTTTGAAAA          | TACATATGAA         | AAGGTAAAAT   | ACTIONACACC           | CUCUS SUCCU                             |
|       |                     |                    | MIGGIANNI    | MCITIACACG            | CIGIARIGGI                              |
| 27701 | ጥርጥልሮልጥርልር          | ראאאראנייאיי       | AAGTTGTGGC   | CCCCACAAAA            | mmomomoca a                             |
|       | ACATGTACTC          | CHANCACIVI         | TTCAACACCG   | CCCCACAAAA            | TIGIGIGGAA                              |
|       | 110110111010        | GIIIGICAIA         | TICANCACCG   | 6666161-1-1-1         | AACACACCTT                              |
| 27751 | AACACTGGCA          | Cutatrate          | CACTGCTATG   | Oma a mm a ca c       | MCOMO COMO                              |
| 21131 | TTCTCACACCCT        | CANACACCAC         | GTGACGATAC   | CTAATTACAG            | TGCTCGCTTT                              |
|       | 11G1GMCCG1          | GWAYGACGAC         | GTGACGATAC   | GATTAATGTC            | ACGAGCGAAA                              |
| 27801 | CCMCMCM3 CC         | Cms cmcms ms       | <b></b>      |                       |   |
| 27601 |                     |                    | TTAAATACAA   |                       |   |
|       | CCAGACATGG          | GATGAGATAT         | AATTTATGTT   | TTCGTCTGCG            | TCGAAATAAC                              |
| 27851 |                     |                    |              |                       |   |
| 2/631 | AGGAAAAGAA          | AATGCCTTAA         | TTTACTAAGT   | TACAAAGCTA            | ATGTCACCAC                              |
|       | TCCTTTCTT           | TTACGGAATT         | AAATGATTCA   | ATGTTTCGAT            | TACAGTGGTG                              |
| 0000  |                     |                    |              | •                     |   |
| 27901 | TAACTGCTTT          | ACTCGCTGCT         | TGCAAAACAA   | ATTCAAAAAG            | TTAGCATTAT                              |
|       | ATTGACGAAA          | TGAGCGACGA         | ACGTTTTGTT   | TAAGTTTTTC            | AATCGTAATA                              |
|       |                     |                    |              |                       |   |
| 27951 | AATTAGA <u>A</u> TA | <u>_GGATTTAAAC</u> | CCCCCGGTCA   | TTTCCTGCTC            | AATACCATTC                              |
|       | TTAATCTTAT          | CCTAAATTTG         | GGGGGCCAGT   | AAAGGACGAG            | TTATGGTAAG                              |
|       |                     |                    |              |                       |   |
| 28001 |                     |                    | GTGGGATATG   |                       |   |
|       | GGGACTTGTT          | AACTGAGATA         | CACCCTATAC   | GAGGTCGCGA            | TGTTGGAACT                              |
|       |                     |                    |              |                       | • "                                     |
| 28051 |                     |                    | AGCATCTGAC   |                       |   |
|       | TCAGTCCGAA          | GGACCTACAG         | TCGTAGACTG   | AAACCGGTCG            | TGGACAGGGC                              |
|       |                     |                    |              |                       |   |
| 28101 | CGGATTTGTT          | CCAGTCCAAC         | TACAGCGACC   | CACCCTAACA            | GAGATGACCA                              |
|       | GCCTAAACAA          | GGTCAGGTTG         | ATGTCGCTGG   | GTGGGATTGT            | CTCTACTGGT                              |
|       |                     |                    |              |                       |   |
| 28151 | ACACAACCAA          | CGCGGCCGCC         | GCTACCGGAC   | TTACATCTAC            | CACAAATACA                              |
|       | TGTGTTGGTT          | GCGCCGGCGG         | CGATGGCCTG   | AATGTAGATG            | GTGTTTATGT                              |
|       |                     |                    |              |                       |   |
| 28201 | CCCCAAGTTT          | CTGCCTTTGT         | CAATAACTGG   | GATAACTTGG            | GCATGTGGTG                              |
|       |                     |                    | GTTATTGACC   |                       |   |
|       |                     |                    |              |                       |   |
| 28251 | GTTCTCCATA          | GCGCTTATGT         | TTGTATGCCT   | TATTATTAT             | TGGCTCATCT                              |
|       |                     |                    |              |                       | ACCGAGTAGA                              |
|       |                     |                    |              |                       |   |
| 28301 | GCTGCCTAAA          | GCGCAAACGC         | GCCCGACCAC   | <b>ርር</b> አዋርጥልጥል ፫   | <b>ጥቦር አጥር አ</b> ጥጥ                     |
|       |                     |                    |              |                       | AGGGTAGTAA                              |
|       |                     |                    |              | CONTRACTOR            | MATURITURE                              |
| 28351 | GTGCTACACC          |                    | ጥርር እ አጥር አጥ | acammon a co          | Carreanara                              |
|       | CACGATGTGG          |                    |              |                       |   |
|       | CUCOUIG100          | GITIGITACI         | ACCI INGGIA  | TOTAMOUTEC            | CIGACTITGT                              |

Figure 27AD

| 28401 | CATGTTCTTT               | TTACAG                    | TATGATTAAA                   | TGAGACATGÁ                  | TTCCTC T                     |
|-------|--------------------------|---------------------------|------------------------------|-----------------------------|------------------------------|
|       | GTACAAGAAA               | AGAGAATGTC                | ATACTAATTT                   | ACTCTGTACT                  | AAGGAGCTCA                   |
| 28451 | TTTTATATTA               | CTGACCCTTG                | TTGCGCTTTT                   | TTGTGCGTGC                  | TCCACATTGG                   |
|       | AAAATATAAT               | GACTGGGAAC                | AACGCGAAAA                   | AACACGCACG                  | AGGTGTAACC                   |
| 28501 | CTGCGGTTTC               | TCACATCGAA                | GTAGACTGCA                   | TTCCAGCCTT                  | CACAGTCTAT                   |
|       | GACGCCAAAG               | AGTGTAGCTT                | CATCTGACGT                   | AAGGTCGGAA                  | GTGTCAGATA                   |
| 28551 | TTGCTTTACG               | GATTTGTCAC                | CCTCACGCTC                   | ATCTGCAGCC                  | TCATCACTGT                   |
|       | AACGAAATGC               | CTAAACAGTG                | GGAGTGCGAG                   | TAGACGTCGG                  | AGTAGTGACA                   |
| 28601 | GGTCATCGCC               | TTTATCCAGT                | GCATTGACTG                   | GGTCTGTGTG                  | CGCTTTGCAT                   |
|       | CCAGTAGCGG               | AAATAGGTCA                | CGTAACTGAC                   | CCAGACACAC                  | GCGAAACGTA                   |
| 28651 | ATCTCAGACA               | CCATCCCCAG                | TACAGGGACA                   | GGACTATAGC                  | TGAGCTTCTT                   |
|       | TAGAGTCTGT               | GCTAGGGGTC                | ATGTCCCTGT                   | CCTGATATCG                  | ACTCGAAGAA                   |
| 28701 | AGAATTCTTT               | AATTATGAAA                | TTTACTGTGA                   | CTTTTCTGCT                  | GATTATTTGC                   |
|       | TCTTAAGAAA               | TTAATACTTT                | AAATGACACT                   | GAAAAGACGA                  | CTAATAAACG                   |
| 28751 | ACCCTATCTG               | CGTTTTGTTC                | CCCGACCTCC                   | AAGCCTCAAA                  | GACATATATC                   |
|       | TGGGATAGAC               | GCAAAACAAG                | GGGCTGGAGG                   | TTCGGAGTTT                  | CTGTATATAG                   |
| 28801 | ATGCAGATTC               | ACTCGTATAT                | GGAATATTCC                   | AAGTTGCTAC                  | AATGAAAAA                    |
|       | TACGTCTAAG               | TGAGCATATA                | CCTTATAAGG                   | TTCAACGATG                  | TTACTTTTTT                   |
| 28851 | GCGATCTTTC               | CGAAGCCTGG                | TTATATGCAA                   | TCATCTCTGT                  | TATGGTGTTC                   |
|       | CGCTAGAAAG               | GCTTCGGACC                | AATATACGTT                   | AGTAGAGACA                  | ATACCACAAG                   |
| 28901 | TCCAGTACCA               | TCTTAGCCCT                | AGCTATATAT                   | CCCTACCTTG                  | ACATTGGCTG                   |
|       | ACGTCATGGT               | AGAATCGGGA                | TCGATATATA                   | GGGATGGAAC                  | TGTAACCGAC                   |
| 28951 | GAACGCAATA               | GATGCCATGA                | ACCACCCAAC                   | TTTCCCCGCG                  | CCCGCTATGC                   |
|       | CTTGCGTTAT               | CTACGGTACT                | TGGTGGGTTG                   | AAAGGGGCGC                  | GGGCGATACG                   |
| 29001 | TTCCACTGCA               | ACAAGTTGTT                | GCCGGCGGCT                   | TTGTCCCAGO                  | CAATCAGCCT                   |
|       | AAGGTGACGT               | TGTTCAACAA                | CGGCCGCCGA                   | AACAGGGTCG                  | GTTAGTCGGA                   |
| 29051 | CGCCCACCTT<br>GCGGGTGGAA | CTCCCACCCC                | CACTGAAATC                   | AGCTACTTTA<br>TCGATGAAA1    | ATCTAACAGG<br>TAGATTGTCC     |
| 29101 | AGGAGATGAC               | TGACACCCTA                | GATCTAGAAA                   | TGGACGGAAT                  | TATTACAGAG                   |
|       | TCCTCTACTG               | ACTGTGGGAT                | CTAGATCTTT                   | ACCTGCCTTA                  | A ATAATGTCTC                 |
| 29151 | CAGCGCCTGC               | TAGAAAGACO                | CAGGGCAGCG                   | GCCGAGCAAC                  | AGCGCATGAA                   |
|       | GTCGCGGACG               | ATCTTTCTGO                | CTCCCGTCGC                   | CGGCTCGTTC                  | CTCGCGTACTT                  |
| 29201 | TCAAGAGCTC<br>AGTTCTCGAG | CAAGACATGO                | TTAACTTGCA<br>AATTGAACGT     | CCAGTGCAA<br>GGTCACGTT      | A AGGGGTATCT<br>T TCCCCATAGA |
| 29251 | TTTGTCTCGT<br>AAACAGAGCA | P AAAGCAGGC               | AAAGTCACCT                   | C ACGACAGTAL<br>A TGCTGTCAT | A TACCACCGGA<br>I ATGGTGGCCT |
| 29301 | CACCGCCTT/<br>GTGGCGGAAT | A GCTACAAGT"<br>CGATGTTCA | r GCCAACCAA(<br>A CGGTTGGTTY | GCAGTCTTT                   | TGGTGGTCAT A ACCACCAGTA      |

Figure 27 AE

| 29351 | GGTGGGAGAA | ACCATTA    | CCATAACTCA | GCACTCGGTA | GAAACC G   |
|-------|------------|------------|------------|------------|------------|
|       | CCACCCTCTT | TTCGGGTAAT | GGTATTGAGT | CGTGAGCCAT | CTTTGGCTTC |
| 29401 | GCTGCATTCA | CTCACCTTGT | CAAGGACCTG | AGGATCTCTG | CACCCTTATT |
|       | CGACGTAAGT | GAGTGGAACA | GTTCCTGGAC | TCCTAGAGAC | GTGGGAATAA |
| 29451 | AAGACCCTGT | GCGGTCTCAA | AGATCTTATT | CCCTTTAACT | AATAAAAAA  |
|       | TTCTGGGACA | CGCCAGAGTT | TCTAGAATAA | GGGAAATTGA | TTATTTTTTT |
| 29501 | AATAATAAAG | CATCACTTAC | TTAAAATCAG | TTAGCAAATT | TCTGTCCAGT |
|       | TTATTATTTC | GTAGTGAATG | AATTTTAGTC | AATCGTTTAA | AGACAGGTCA |
| 29551 | TTATTCAGCA | GCACCTCCTT | GCCCTCCTCC | CAGCTCTGGT | ATTGCAGCTT |
|       | AATAAGTCGT | CGTGGAGGAA | CGGGAGGAGG | GTCGAGACCA | TAACGTCGAA |
| 29601 | CCTCCTGGCT | GCAAACTTTC | TCCACAATCT | AAATGGAATG | TCAGTTTCCT |
|       | GGAGGACCGA | CGTTTGAAAG | AGGTGTTAGA | TTTACCTTAC | AGTCAAAGGA |
| 29651 |            | TCCATCCGCA |            |            |            |
|       | GGACAAGGAC | AGGTAGGCGT | GGGTGATAGA | AGTACAACAA | CGTCTACTTC |
| 29701 |            | CGTCTGAAGA |            |            |            |
|       |            | GCAGACTTCT |            |            |            |
| 29751 |            | CCTCCAACTG |            |            |            |
|       | CCTTTGGCCA | GGAGGTTGAC | ACGGAAAAGA | ATGAGGAGGG | AAACATAGGG |
| 29801 |            | TCAAGAGAGT |            |            |            |
|       |            | AGTTCTCTCA |            |            |            |
| 29851 | GAACCTCTAG | TTACCTCCAA | TGGCATGCTT | GCGCTCAAAA | TGGGCAACGG |
|       |            | AATGGAGGTT |            |            |            |
| 29901 |            | GACGAGGCCG |            |            |            |
|       |            | CTGCTCCGGC |            | -          |            |
| 29951 |            | TCTCAAAAA  |            |            |            |
|       |            | AGAGTTTTTT |            |            |            |
| 30001 | GCACCCCTCA | CAGTTACCTC | AGAAGCCCTA | ACTGTGGCTG | CCGCCGCACC |
|       |            | GTCAATGGAG |            |            |            |
| 30051 |            | GCGGGCAACA |            |            |            |
|       |            | CGCCCGTTGT |            |            |            |
| 30101 | CCGTGCACGA |            |            |            |            |
|       | GGCACGTGCT |            |            |            |            |
| 30151 | TCAGAAGGAA |            |            |            |            |
|       | AGTCTTCCTT |            |            |            |            |
| 30201 | TAGCAGTACC |            |            |            |            |
|       | ATCGTCATGG |            |            |            |            |
| 30251 | GTAGCTTGGG | CATTGACTTG | AAAGAGCCCA | TTTATACACA | AAATGGAAAA |
|       | CATCGAACCC | GTAACTGAAC | TTTCTCGGGT | AAATATGTGT | TTTACCTTTT |

Figure 27 AF

| 30301 | CTAGGACTAA               | A CGGGGC                 | TCCTTTGCAT               | GTAACAGAČG               | ACCIAN                   |
|-------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
|       |                          | Temecccc                 |                          |                          |                          |
| 30351 | TTTGACCGTA<br>AAACTGGCAT | GCAACTGGTC<br>CGTTGACCAG |                          |                          |                          |
| 30401 | AAACTAAAGT               | TACTGGAGCC               | TTGGGTTTTG               | ATTCACAAGG               | CAATATGCAA               |
|       | TTTGATTTCA               | ATGACCTCGG               | AACCCAAAAC               | TAAGTGTTCC               | GTTATACGTT               |
| 30451 | CTTAATGTAG               | CAGGAGGACT<br>GTCCTCCTGA |                          |                          |                          |
| 20501 |                          | AGTTATCCGT               |                          |                          |                          |
| 30501 | TGAACTACAA               |                          |                          |                          |                          |
| 30551 |                          | CCCTCTTTTT               |                          |                          |                          |
|       |                          | GGGAGAAAA                |                          |                          |                          |
| 30601 |                          | GCCTTTACTT<br>CGGAAATGAA |                          |                          |                          |
| 30651 |                          | CTAAGCACTG               |                          |                          |                          |
|       | ACTCCAATTG               | GATTCGTGAC               | GGTTCCCCAA               | CTACAAACTG               | CGATGTCGGT               |
| 30701 |                          | TGCAGGAGAT<br>ACGTCCTCTA |                          |                          | TAATGCACCA               |
|       |                          |                          |                          |                          | •                        |
| 30751 |                          | CCCTCAAAAC<br>GGGAGTTTTG |                          |                          |                          |
| 30801 |                          |                          |                          |                          | TTTGACAGCA               |
|       |                          |                          |                          |                          | AAACTGTCGT               |
| 30851 |                          | TACAGTAGGA<br>ATGTCATCCT |                          |                          | AACTTTGTGG               |
|       |                          |                          |                          |                          |                          |
| 30901 |                          | CTCCATCTCC<br>GAGGTAGAGG |                          |                          | AGAAAGATGC               |
| 30951 | <b>ጥአአአርጥር ኦር</b> ጥ      | ጥጥር-ርጥር ጥጥ <u>አ</u>      | САДАТСТСС                | СРСТСАВАТА               | CTTGCTACAG               |
| 20321 |                          |                          |                          |                          | GAACGATGTC               |
| 31001 |                          |                          |                          |                          | TGGAACAGTT               |
|       |                          |                          |                          |                          | ACCTTGTCAA               |
| 31051 | CAAAGTGCTC               | ATCTTATTAT<br>TAGAATAATA | AAGATTTGAC<br>TTCTAAACTG | GAAAATGGAG<br>CTTTTACCTC | TGCTACTAAA<br>ACGATGATTT |
| 21101 |                          |                          |                          |                          | GGAGATCTTA               |
| 31101 | GTTAAGGAAG               | GACCTGGGTC               | TTATAACCTT               | GAAATCTTTA               | CCTCTAGAAT               |
| 31151 | CTGAAGGCAC               | AGCCTATACA               | AACGCTGTTG               | GATTTATGCC               | TAACCTATCA               |
|       |                          |                          |                          |                          | ATTGGATAGT               |
| 31201 | GCTTATCCAA<br>CGAATAGGTT | AATCTCACGG<br>TTAGAGTGCC | TAAAACTGCC<br>ATTTTGACGG | AAAAGTAACA<br>TTTTCATTGT | TTGTCAGTCA<br>AACAGTCAGT |

Figure 27 AG

| 31251  |   | AMEGGAGACA<br>TTGCCTCTGT |      |                          |
|--------|---|--------------------------|------|--------------------------|
| 31301  |   | ACAGGAAACA<br>TGTCCTTTGT |      |                          |
| 31351  |   | GGGACTGGTC<br>CCCTGACCAG | <br> |                          |
| 31401  |   | TACACTTTTT<br>ATGTGAAAAA | <br> | ***********              |
| 31451  |   | TCAACGTGTT<br>AGTTGÇACAA | <br> |                          |
| 31501  | TTTTTCATTC<br>AAAAAGTAAG                | AGTAGTATAG<br>TCATCATATC | <br> |                          |
| 31551  |   | TCAAACTCAC<br>AGTTTGAGTG |      |                          |
| 31601  |   | CAGAGTACAC<br>GTCTCATGTG |      |                          |
| 31:651 | *************************************** | TGGGTAACAG<br>ACCCATTGTC | <br> |                          |
| 31701  |   | AGCCAAACGC<br>TCGGTTTGCG |      |                          |
| 31751  |   | AGTTCATGTC<br>TCAAGTACAG |      | CAGGCTGCTG<br>GTCCGACGAC |
| 31801  |   | GGTTGCTTAA<br>CCAACGAATT | <br> |                          |
| 31851  |   | GTCATAATCG<br>CAGTATTAGC | <br> |                          |
| 31901  |   | TAAACTGCTG<br>ATTTGACGAC | <br> |                          |
| 31951  |   | GTCTCCTCAG<br>CAGAGGAGTC |      |                          |
| 32001  |   |                          |      | TAAATCAGCA<br>ATTTAGTCGT |
| 32051  |   |                          | <br> | CACAGTGCAA<br>GTGTCACGTT |
| 32101  |   |                          |      | acgtggccat<br>tgcaccggta |
|        | CATACCACAA<br>GTATGGTGTT                |                          |      | AAACACGCTG<br>TTTGTGCGAC |

Figure 27 AH

| 32201  | GACATAAACA                    | CTCTTT                     | TGGCATGTTG            | TAATTCACCA                     | CCTCCCTA   |
|--------|-------------------------------|----------------------------|-----------------------|--------------------------------|------------|
|        | CTGTATTTGT                    | AATGGAGAAA                 | ACCGTACAAC            | ATTAAGTGGT                     | GGAGGGCCAT |
| 32251  | CCATATAAAC                    | CTCTGATTAA                 | ACATGGCGCC            | ATCCACCACC                     | ATCCTAAACC |
|        |                               | GAGACTAATT                 |                       |                                |            |
| 32301  |                               | AACCTGCCCG                 |                       |                                |            |
|        | TCGACCGGTT                    | TTGGACGGGC                 | GGCCGATATG            | TGACGTCCCT                     | TGGCCCTGAC |
| 32351  | GAACAATGAC                    | AGTGGAGAGC                 | CCAGGACTCG            | TAACCATGGA                     | TCATCATGCT |
|        | CTTGTTACTG                    | TCACCTCTCG                 | GGTCCTGAGC            | ATTGGTACCT                     | AGTAGTACGA |
| 32401  | CGTCATGATA                    | TCAATGTTGG                 | CACAACACAG            | GCACACGTGC                     | ATACACTTCC |
|        | GCAGTACTAT                    | AGTTACAACC                 | GTGTTGTGTC            | CGTGTGCACG                     | TATGTGAAGG |
| 32451  | TCAGGATTAC                    | AAGCTCCTCC                 | CGCGTTAGAA            | CCATATCCCA                     | GGGAACAACC |
|        |                               | TTCGAGGAGG                 |                       |                                |            |
| 32501  | CATTCCTGAA                    | TCAGCGTAAA                 | TCCCACACTG            | CAGGGAAGAC                     | CTCGCACGTA |
| 32302  |                               | AGTCGCATTT                 |                       |                                |            |
| 32551  | <b>ል</b> ሮጥሮል ሮርጥጥር           | TGCATTGTCA                 | AAGTGTTACA            | TTCGGGCAGC                     | AGCGGATGAT |
| 32332  |                               | ACGTAACAGT                 |                       |                                |            |
| 32601  | ССТССАСТАТ                    | GGTAGCGCGG                 | GTTTCTGTCT            | CAAAAGGAGG                     | TAGACGATCC |
| 52,001 |                               | CCATCGCGCC                 |                       |                                |            |
| 32651  | СТАСТСТАСС                    | GACTECCCC                  | AGACAACCGA            | GATCGTGTTG                     | GTCGTAGTGT |
| 52051  |                               |                            |                       |                                | CAGCATCACA |
| 32701  | CATCCCAAAT                    | GGAACGCCGG                 | ACGTAGTCAT            | ATTTCCTGAA                     | GCAAAACCAG |
|        | GTACGGTTTA                    | CCTTGCGGCC                 | TGCATCAGTA            | TAAAGGACTT                     | CGTTTTGGTC |
| 32751  | CTCCCCCCT                     | GACAAACAGA                 | TCTGCGTCTC            | CGGTCTCGCC                     | GCTTAGATCG |
| 32/31  | CACGCCCGCA                    | CTGTTTGTCT                 | AGACGCAGAG            | GCCAGAGCGG                     | CGAATCTAGC |
|        |                               |                            |                       |                                |            |
| 32801  |                               | TAGTTGTAGT                 |                       |                                |            |
|        | GAGACACATC                    | ATCAACATCA                 | TATAGGTGAG            | AGAGTTTCGT                     | AGGTCCGCGG |
| 32851  |                               |                            |                       |                                | TGCCCTGATA |
|        | GGGACCGAAG                    | CCCAAGATAC                 | ATTTGAGGAA            | GTACGCGGCG                     | ACGGGACTAT |
| 32901  | ACATCCACCA                    | CCGCAGAATA                 | AGCCACACCC            | AGCCAACCTA                     | CACATTCGTT |
|        | TGTAGGTGGT                    | GGCGTCTTAT                 | TCGGTGTGGG            | TCGGTTGGAT                     | GTGTAAGCAA |
| 32951  | CTGCGAGTCA                    | CACACGGGAG                 | GAGCGGGAAG            | AGCTGGAAGA                     | ACCATGTTTT |
| 02202  | GACGCTCAGT                    | GTGTGCCCTC                 | CTCGCCCTTC            | TCGACCTTCT                     | TGGTACAAAA |
| 33001  | بابعله الإمليمانية والمتاسمات | ССВВВЕЕТ                   | ATCCAAAACC            | TCAAAATGAA                     | GATCTATTAA |
| 33001  | AAAAAAAA                      | GGTTTTCTAA                 | TAGGTTTTGG            | AGTTTTACTT                     | CTAGATAATT |
| 33051  | こかごり かいししい                    | <b>がたしてかんしい</b>            | <b>ጥር</b> ርርርጥርርጥ     | <u>አአአሮ</u> ፓርፕልቦ <sup>ኢ</sup> | GCCAAAGAAC |
| 22021  | CACTTGCGCG                    | AGGGGAGGCC                 | ACCGCACCAG            | TTTGAGATGT                     | CGGTTTCTTG |
| 22101  | አርንሞአ <b>አጥር</b> ር            | ע בו עי עיני אוריי אוריי ע | <b>ጥርጥጥር</b> ር ልር ል ል | ብ <b>ረ</b> ርርርምፕሮር እ           | AAGGCAAACG |
| 33101  | TCTATTACCG                    | TAAACATTCT                 | ACAACGTGTT            | ACCGAAGGTT                     | TTCCGTTTGC |

Figure 27 AI

| 33151 | GCCCTCACGT<br>CGGGAGTGCA | GTGGAC<br>GTCACCTG       | GTAAAGGCTA<br>CATTTCCGAT | AACCCTTCAG<br>TTGGGAAGTC | TGTGAAT TC               |
|-------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| 33201 | CTCTATAAAC               | ATTCCAGCAC               | CTTCAACCAT               | GCCCAAATAA               | TTCTCATCTC               |
| 33251 |                          | TAAGGTCGTG               |                          |                          |                          |
|       |                          | GTTATATAGA               |                          |                          | •                        |
| 33301 |                          | TCTGCTCCAG<br>AGACGAGGTC |                          |                          |                          |
| 33351 |                          | GCAAAAATTC<br>CGTTTTTAAG |                          |                          |                          |
| 33401 |                          | TTAACAAAAA<br>AATTGTTTTT |                          |                          |                          |
| 33451 |                          | ATAATCGTGC<br>TATTAGCACG |                          |                          |                          |
| 33501 |                          | CCATGACAAA<br>GGTACTGTTT |                          |                          |                          |
| 33551 |                          | CTAACCAGCG<br>GATTGGTCGC |                          |                          |                          |
| 33601 |                          | ATGCAAGGTG<br>TACGTTCCAC |                          |                          |                          |
| 33651 |                          | GCACATCGTA<br>CGTGTAGCAT |                          |                          |                          |
| 33701 |                          | ACCACAGAAA<br>TGGTGTCTTT |                          |                          |                          |
| 33751 |                          | CATAAACACA<br>GTATTTGTGT |                          |                          |                          |
| 3380i |                          | TCTTACAACA<br>AGAATGTTGT |                          |                          | -                        |
| 33851 |                          | TGCCGGCGTG<br>ACGGCCGCAC |                          |                          |                          |
| 33901 |                          |                          |                          |                          | ATGTAAGACT<br>TACATTCTGA |
| 33951 |                          | -                        |                          |                          | AAAGCGACCG<br>TTTCGCTGGC |
| 34001 |                          |                          |                          |                          | ACATTACAGC<br>TGTAATGTCG |
| 34051 |                          |                          |                          |                          | ACATAAACAC<br>TGTATTTGTG |

Figure 27AJ

| 34101   | CTGAAAAACC I<br>GACTTTTTGG   | CPSTTGCCTA<br>G. ACGGAT   | GGCAAAATAG<br>CCGTTTTATC   | CACCCTCC&G   | CAGGTC T   |
|---|--|---|--|--|--|
| 34151   | ACATACAGCG<br>TGTATGTCGC   | CTTCCACAGC<br>GAAGGTGTCG  | GGCAGCCATA<br>CCGTCGGTAT   | ACAGTCAGCC<br>TGTCAGTCGG   | TTACCAGTAA<br>AATGGTCATT   |
| 34201   | AAAAGAAAAC<br>TTTTCTTTTG   | CTATTAAAAA<br>GATAATTTT   | AACACCACTC<br>TTGTGGTGAG   | GACACGGCAC<br>CTGTGCCGTG   | CAGCTCAATC<br>GTCGAGTTAG   |
| 34251   | AGTCACAGTG<br>TCAGTGTCAC   |   |  |  |  |
| 34301   | AAAAATGACG<br>TTTTTACTGC   |   |  | AAACACCCAG<br>TTTGTGGGTC   |  |
| 34351   | GCGAACCTAC<br>CGCTTGGATG   | GCCCAGAAAC<br>CGGGTCTTTG  | GAAAGCCAAA<br>CTTTCGGTTT   | AAACCCACAA<br>TTTGGGTGTT   | CTTCCTCAAA<br>GAAGGAGTTT   |
| 34401   |  |   |  | CTTCCCATTT<br>GAAGGGTAAA   |  |
| 34451   | ACAATTCCCA<br>TGTTAAGGGT   | ACACATACAA<br>TGTGTATGTT  | GTTACTCCGC<br>CAATGAGGCG   | CCTAAAACCT<br>GGATTTTGGA   | ACGTCACCCG<br>TGCAGTGGGC   |
| 34501   | CCCCGTTCCC<br>GGGGCAAGGG   | ACGCCCCCGC<br>TGCGGGGGGCGC  | CCACGTCACA<br>GGTGCAGTGT   | AACTCCACCC<br>TTGAGGTGGG   | CCTCATTATC<br>GGAGTAATAG   |
|   |  |   |  |  | PacI   |
|   |  |   |  |  |  |
| 34551   |  |   |  | ТАТТGАТGАТ<br>АТААСТАСТА   |  |
| 34551<br>34601  | TATAACCGAA   | GTTAGGTTTT<br>TGCGACGCGA  | ATTCCATATA   | ATAACTACTA   |  |
|   | TATAACCGAA  AATTCGGATC TTAAGCCTAG  CTCGCTTCCG  | GTTAGGTTTT  TGCGACGCGA  ACGCTGCGCT  GCGGCATCGG  | ATTCCATATA GGCTGGATGG CCGACCTACC GATGCCCGCG  | ATAACTACTA<br>CCTTCCCCAT<br>GGAAGGGGTA   | CAATTAATTC TATGATTCTT ATACTAAGAA TGCTGTCCAG  |
| 34601   | TATAACCGAA  AATTCGGATC TTAAGCCTAG  CTCGCTTCCG GAGCGAAGGC GCAGGTAGAT  | GTTAGGTTTT TGCGACGCGA ACGCTGCGCT GCGGCATCGG CGCCGTAGCC  | ATTCCATATA GGCTGGATGG CCGACCTACC GATGCCCGCG CTACGGGCGC AGGGACAGCT  | ATAACTACTA CCTTCCCCAT GGAAGGGGTA TTGCAGGCCA AACGTCCGGT TCAAGGCCAG  | CAATTAATTC TATGATTCTT ATACTAAGAA TGCTGTCCAG  |
| 34601   | TATAACCGAA  AATTCGGATC TTAAGCCTAG  CTCGCTTCCG GAGCGAAGGC GCAGGTAGAT CGTCCATCTA  GGAACCGTAA   | TGCGACGCGA<br>ACGCTGCGCT<br>GCGGCATCGG<br>CGCCGTAGCC<br>GACGACCATC<br>CTGCTGGTAG  | ATTCCATATA GGCTGGATGG CCGACCTACC GATGCCCGCG CTACGGGCGC AGGGACAGCT TCCCTGTCGA   | ATAACTACTA  CCTTCCCAT GGAAGGGTA  TTGCAGGCCA AACGTCCGGT  TCAAGGCCAG AGTTCCGGTC  TTTTCCATAG  | TATGATTCTT ATACTAAGAA TGCTGTCCAG ACGACAGGTC CAAAAGGCCA   |
| 34601<br>34651<br>34701<br>34751                            | TATAACCGAA  AATTCGGATC TTAAGCCTAG  CTCGCTTCCG GAGCGAAGGC GCAGGTAGAT CGTCCATCTA GGAACCGTAA CCTTGGCATT   | TTAGGTTTT TGCGACGCGA ACGCTGCGCT GCGCCATCGG CGCCGTAGCC GACGACCATC CTGCTGGTAG AAAGGCCGCG TTTCCGGCGC   | ATTCCATATA GGCTGGATGC CCGACCTACC GATGCCCGCG CTACGGGCGC AGGGACAGCT TCCCTGTCGA TTGCTGGCGT AACGACCGCA   | ATAACTACTA  CCTTCCCCAT GGAAGGGGTA  TTGCAGGCCA AACGTCCGGT  TCAAGGCCAG AGTTCCGGTC  TTTTCCATAG AAAAGGTATC AGTCAGAGGT  | CAATTAATTC TATGATTCTT ATACTAAGAA TGCTGTCCAG ACGACAGGTC CAAAAGGCCA GTTTTCCGGT GCTCCGCCCC CGAGGCGGGG                                       |
| 34601<br>34651<br>34701<br>34751<br>34801                   | TATAACCGAA  AATTCGGATC TTAAGCCTAG  CTCGCTTCCG GAGCGAAGGC GCAGGTAGAT CGTCCATCTA  GGAACCGTAA CCTTGGCATT  CCTGACGAGC GGACTGCTCG   | GTTAGGTTTT TGCGACGCGA ACGCTGCGCT GCGCCATCGG CGCCGTAGCC GACGACCATC CTGCTGGTAG AAAGGCCGCG TTTCCGGCGC ATCACAAAAA TAGTGTTTTT                                  | ATTCCATATA GGCTGGATGG CCGACCTACC GATGCCCGCG CTACGGGCGC AGGGACAGCT TCCCTGTCGA TTGCTGGCGT AACGACCGCA TCGACGCTCA AGCTGCGAGT                               | ATAACTACTA CCTTCCCCAT GGAAGGGGTA TTGCAGGCCA AACGTCCGGT TCAAGGCCAG AGTTCCATAG AAAAGGTATC AGTCAGAGGT TCAGTCTCCA  | TATGATTCTT ATACTAAGAA TGCTGTCCAG ACGACAGGTC CAAAAGGCCA GTTTTCCGGT GCTCCGCCCC CGAGGCGGGG  |
| 34501<br>34551<br>34701<br>34751<br>34801<br>34851          | TATAACCGAA  AATTCGGATC TTAAGCCTAG  CTCGCTTCCG GAGCGAAGGC GCAGGTAGAT CGTCCATCTA CCTTGGCATT CCTGACGAGC GGACTGCTCG GACAGGACTA CTTGCCTTCG GACAGGACTA CTGTCCTGAT  | GTTAGGTTTT TGCGACGCGA ACGCTGCGCT GCGGCATCGG CGCCGTAGCC GACGACCATC CTGCTGGTAG AAAGGCCGCG ATCACAAAAA TAGTGTTTTT TAAAGATACC ATTTCTATGG                       | ATTCCATATA  GGCTGGATGG CCGACCTACC  GATGCCCGCG CTACGGGCGC  AGGGACAGCT TCCCTGTCGA TTGCTGGCGT AACGACGCTCA AGCTGCGAGT AGCTGCGAGT CCGCAAAGG                 | ATAACTACTA CCTTCCCAT GGAAGGGGTA TTGCAGGCCA AACGTCCGGT TCAAGGCCAG AGTTCCATAG AAAAGGTATC AGTCAGAGGT TCAGTCTCCA CCCTGGAAGC GGGACCTTCG                                   | TATGATTCTT ATACTAAGAA  TGCTGTCCAG ACGACAGGTC  CAAAAGGCCA GTTTTCCGGT  GCTCCGCCCC CGAGGCGGGG GCGAAACCC CCGCTTTGGG TCCCTCGTGC               |
| 34501<br>34651<br>34701<br>34751<br>34801<br>34851<br>34901 | TATAACCGAA  AATTCGGATC TTAAGCCTAG  CTCGCTTCCG GAGCGAAGGC  GCAGGTAGAT CGTCCATCTA  CCTTGGCATT  CCTGACGAGC GGACTGCTCG GACAGGACTA  CTGTCCTGAT  CTGTCCTGAT  CTGTCCTGAT  CTGTCCTGAT  CCTCCTGGTAA  CCTTCCGGAACACACACACACACACACACACACACACACA | GTTAGGTTTT TGCGACGCGA ACGCTGCGCT GCGCCATCGG CGCCGTAGCC GACGACCATC CTGCTGGTAG AAAGGCCGCG ATCACAAAAA TAGTGTTTTT TAAAGATACC ATTTCTATGG TCCGACCCTG AGGCTGGGAC | ATTCCATATA GGCTGGATGG CCGACCTACC GATGCCCGCG CTACGGGCGC AGGGACAGCT TCCCTGTCGA TTGCTGGCGT AACGACCGCA AGGCTTCC AGGCGTTTCC CGCCAAAGG CCGCTTACCG GGCGAATGGC | ATAACTACTA CCTTCCCCAT GGAAGGGGTA TTGCAGGCCA AACGTCCGGT TCAAGGCCAG AGTTCCGGTC TTTTCCATAG AAAAGGTATC AGTCAGAGGT TCAGTCTCCA CCCTGGAAGC GGGACCTTCG GATACCTGTC CTATGGACAG | TATGATTCTT ATACTAAGAA  TGCTGTCCAG ACGACAGGTC  CAAAAGGCCA GTTTTCCGGT  GCTCCGCCCC CGAGGCGGGG  GCCGAAACCC CCGCTTTGGG  TCCCTCGTGC AGGGAGCACG |

Figure 27 AK

| 35051 |   | CCGCTGCGCC<br>GGCGACGCGG |      |                          |
|-------|---|--------------------------|------|--------------------------|
| 35101 |   | ACGACTTATC<br>TGCTGAATAG |      |                          |
| 35151 |   | AGGTATGTAG<br>TCCATACATC |      |                          |
| 35201 |   | CTACACTAGA<br>GATGTGATCT |      |                          |
| 35251 |   | CCTTCGGAAA<br>GGAAGCCTTT | •    |                          |
| 35301 |   | GGTAGCGGTG<br>CCATCGCCAC | <br> |                          |
| 35351 | • | AGGATCTCAA<br>TCCTAGAGTT | <br> |                          |
| 35401 |   | GGAACGAAAA<br>CCTTGCTTTT |      |                          |
| 35451 |   | ATCTTCACCT<br>TAGAAGTGGA | <br> |                          |
| 35501 |   | TGGTCTGACA<br>ACCAGACTGT | <br> | GAGGCACCTA<br>CTCCGTGGAT |
| 35551 |   | CTGTCTATTT<br>GACAGATAAA |      |                          |
| 35601 |   | CTACGATACG<br>GATGCTATGC | <br> |                          |
| 35651 |   | CGAGACCCAC<br>GCTCTGGGTG |      |                          |
| 35701 |   | CGGAAGGGCC<br>GCCTTCCCGG | <br> |                          |
| 35751 |   | AGTCTATTAA<br>TCAGATAATT |      | TAAGTAGTTC<br>ATTCATCAAG |
| 35801 |   |                          | <br> | GGCATCGTGG<br>CCGTAGCACC |
| 35851 |   |                          |      | TTCCCAACGA<br>AAGGGTTGCT |
| 35901 |   |                          |      | CGGTTAGCTC<br>GCCAATCGAG |
| 35951 |   |                          | <br> | GTGTTATCAC<br>CACAATAGTG |

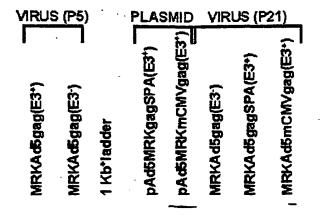
Figure 27AL

| WO 02/2 | 2080                     |                          |                          |                          |                         | PCT/US01/28861 |
|---------|--------------------------|--------------------------|--------------------------|--------------------------|-------------------------|----------------|
| 36001   | TCATGGTTAT AGTACCAATA    | AGCACTG<br>CCGTCGTGAC    | CATAATTCTC<br>GTATTAAGAG | TTACTGTCAT<br>AATGACAGTA | GCCATC TA               | <u>.</u>       |
| 36051   | AGATGCTTTT<br>TCTACGAAAA |                          | TGAGTACTCA<br>ACTCATGAGT |                          |                         |                |
| 36101   | GTGTATGCGG<br>CACATACGCC |                          | GCTCTTGCCC<br>CGAGAACGGG |                          |                         |                |
| 36151   | CCGCGCCACA<br>GGCGCGGTGT | TAGCAGAACT<br>ATCGTCTTGA | TTAAAAGTGC<br>AATTTTCACG | TCATCATTGG<br>AGTAGTAACC | AAAACGTTC'<br>TTTTGCAAG | r<br>A         |
| 36201   | TCGGGGCGAA<br>AGCCCCGCTT |                          | GATCTTACCG<br>CTAGAATGGC |                          |                         |                |
| 36251   | GTAACCCACT<br>CATTGGGTGA |                          | ACTGATCTTC<br>TGACTAGAAG |                          |                         |                |
| 36301 - | GCGTTTCTGG<br>CGCAAAGACC | GTGAGCAAAA<br>CACTCGTTTT | ACAGGAAGGC<br>TGTCCTTCCG | AAAATGCCGC<br>TTTTACGGCG | AAAAAAGGG<br>TTTTTTCCC  | A<br>P         |
| 36351   | ATAAGGGCGA<br>TATTCCCGCT | CACGGAAATG<br>GTGCCTTTAC | TTGAATACTC<br>AACTTATGAG | ATACTCTTCC<br>TATGAGAAGG | TTTTTCAAT.              | A<br>T         |
| 36401   | TTATTGAAGC<br>AATAACTTCG |                          | GTTATTGTCT<br>CAATAACAGA |                          |                         |                |
| 36451   | AATGTATTTA<br>TTACATAAAT |                          | CAAATAGGGG<br>GTTTATCCCC |                          |                         |                |
| 36501   | AAAGTGCCAC<br>TTTCACGGTG |                          | AGAAACCATT<br>TCTTTGGTAA |                          |                         |                |
| 36551   | TAAAAATAGG<br>ATTTTTATCC |                          | GGCCCTTTCG<br>CCGGGAAAGC |                          |                         |                |
|         |                          | PacI                     |                          |                          |                         |                |

36601 ATTCTTAATT TCTTAATTAA (SEQ ID NO:34)

TAAGAATTAA AGAATTAATT (SEQ ID NO:35)

Figure 27AM



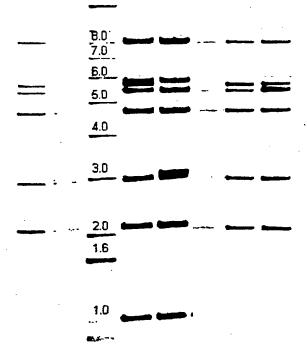


FIGURE 28

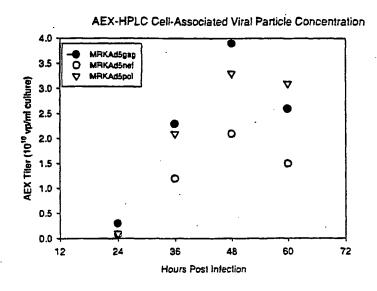


FIGURE 29A

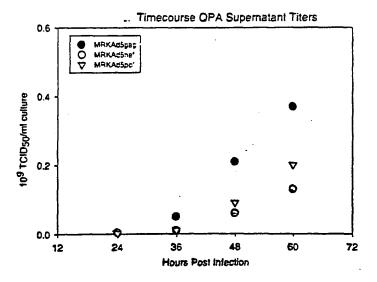


FIGURE 29B

| atg<br>Met<br>1   | gat<br>Asp | gca<br>Ala | atg<br>Met        | aag<br>Lys<br>5   | aga<br>Arg        | GJÀ<br>āāā | ctc<br>Leu | tgc<br>Cys | tgt<br>Cys<br>10  | gtg<br>Val        | ctg<br>Leu | ctg<br>Leu | ctg<br>Leu | tgt<br>Cys<br>15  | gga<br>Gly        | 48         |
|-------------------|------------|------------|-------------------|-------------------|-------------------|------------|------------|------------|-------------------|-------------------|------------|------------|------------|-------------------|-------------------|------------|
|                   |            |            | gtt<br>Val<br>20  |                   |                   |            |            |            |                   |                   |            |            |            |                   |                   | 96         |
|                   |            |            | agg<br>Arg        |                   |                   |            |            |            |                   |                   |            |            |            |                   |                   | 144        |
|                   |            |            | cag<br>Gln        |                   |                   |            |            |            |                   |                   |            |            |            |                   |                   | 192        |
| tct<br>Ser<br>65  | gag<br>Glu | gag<br>Glu | ctg<br>Leu        | agg<br>Arg        | tcc<br>Ser<br>70  | ctg<br>Leu | tac<br>Tyr | aac<br>Asn | aca<br>Thr        | gtg<br>Val<br>75  | gct<br>Ala | acc<br>Thr | ctg<br>Leu | tac<br>Tyr        | tgt<br>Cys<br>80  | 240        |
|                   |            |            | aag<br>Lys        |                   |                   |            |            |            |                   |                   |            |            |            |                   |                   | 288        |
|                   |            |            | gag<br>Glu<br>100 |                   |                   |            |            |            |                   |                   |            |            |            |                   |                   | 336        |
|                   |            |            | ggc<br>Gly        |                   |                   |            |            |            |                   |                   |            |            |            |                   |                   | 384        |
|                   |            |            | cag<br>Gln        |                   |                   |            |            |            |                   |                   |            |            |            |                   |                   | 432        |
|                   |            |            | tgg<br>Trp        |                   |                   |            |            |            |                   |                   |            |            |            |                   |                   | 480        |
|                   |            |            | atg<br>Met        |                   |                   |            |            |            |                   |                   |            |            |            |                   |                   | 528        |
|                   |            |            | atg<br>Met<br>180 |                   |                   |            |            |            |                   |                   |            |            |            |                   |                   | 576        |
|                   |            |            | gag<br>Glu        |                   |                   |            |            |            |                   |                   |            |            |            |                   |                   | 624        |
|                   |            |            | cac<br>His        |                   |                   |            |            |            |                   |                   |            |            |            |                   |                   | 672        |
| agg<br>Arg<br>225 | ggc<br>Gly | tct<br>Ser | gac<br>Asp        | att<br>Ile        | gct<br>Ala<br>230 | ggc<br>Gly | acc<br>Thr | acc<br>Thr | tcc<br>Ser        | acc<br>Thr<br>235 | ctc<br>Leu | cag<br>Gln | gag<br>Glu | cag<br>Gln        | att<br>Ile<br>240 | <b>720</b> |
| ggc<br>Gly        | tgg<br>Trp | atg<br>Met | acc<br>Thr        | aac<br>Asn<br>245 | aac<br>Asn        | ccc<br>Pro | ccc<br>Pro | atc<br>Ile | cct<br>Pro<br>250 | gtg<br>Val        | Gly        | gaa<br>Glu | atc<br>Ile | tac<br>Tyr<br>255 | aag<br>Lys        | 768        |

Figure 30'A"

| agg<br>Arg        | tgg<br>Trp        | atc<br>Ile        | atc<br>Ile<br>260 | ctg<br>Leu        | Gly<br>ggc        | ctg<br>Leu         | aac<br>Asn        | aag<br>Lys<br>265 | att<br>Ile        | gtg<br>Val             | agg<br>Arg        | atg<br>Met        | tac<br>Tyr<br>270 | tcc<br>Ser        | ccc<br>Pro        | 816  |
|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|--------------------|-------------------|-------------------|-------------------|------------------------|-------------------|-------------------|-------------------|-------------------|-------------------|------|
| acc<br>Thr        | tcc<br>Ser        | atc<br>Ile<br>275 | ctg<br>Leu        | gac<br>Asp        | atc<br>Ile        | agg<br><b>A</b> rg | cag<br>Gln<br>280 | Gly               | ccc<br>Pro        | aag<br>Lys             | gag<br>Glu        | ccc<br>Pro<br>285 | ttc<br>Phe        | agg<br>Arg        | gac<br>Asp        | 864  |
| tat<br>Tyr        | gtg<br>Val<br>290 | gac<br>Asp        | agg<br>Arg        | ttc<br>Phe        | tac<br>Tyr        | aag<br>Lys<br>295  | acc<br>Thr        | ctg<br>Leu        | agg<br>Arg        | gct<br>Ala             | gag<br>Glu<br>300 | cag<br>Gln        | gcc<br>Ala        | tcc<br>Ser        | cag<br>Gln        | 912  |
| gag<br>Glu<br>305 | gtg<br>Val        | aag<br>Lys        | aac<br>Asn        | tgg<br>Trp        | atg<br>Met<br>310 | aca<br>Thr         | gag<br>Glu        | acc<br>Thr        | ctg<br>Leu        | ctg<br>Leu<br>315      | gtg<br>Val        | cag<br>Gln        | aat<br>Asn        | gcc<br>Ala        | aac<br>Asn<br>320 | 960  |
| cct<br>Pro        | gac<br>Asp        | tgc<br>Cys        | aag<br>Lys        | acc<br>Thr<br>325 | atc<br>Ile        | ctg<br>Leu         | aag<br>Lys        | gcc<br>Ala        | ctg<br>Leu<br>330 | ggc<br>Gly             | cct<br>Pro        | gct<br>Ala        | gcc<br>Ala        | acc<br>Thr<br>335 | ctg<br>Leu        | 1008 |
| gag<br>Glu        | gag<br>Glu        | atg<br>Met        | atg<br>Met<br>340 | aca<br>Thr        | gcc<br>Ala        | tgc<br>Cys         | cag<br>Gln        | ggg<br>Gly<br>345 | gtg<br>Val        | Gj <sup>y</sup><br>ggg | ggc<br>Gly        | cct<br>Pro        | ggt<br>Gly<br>350 | cac<br>His        | aag<br>Lys        | 1056 |
| gcc<br>Ala        | agg<br>Arg        | gtg<br>Val<br>355 | ctg<br>Leu        | gct<br>Ala        | gag<br>Glu        | gcc<br>Ala         | atg<br>Met<br>360 | tcc<br>Ser        | cag<br>Gln        | gtg<br>Val             | acc<br>Thr        | aac<br>Asn<br>365 | tcc<br>Ser        | gcc<br>Ala        | acc<br>Thr        | 1104 |
| atc<br>Ile        | atg<br>Met<br>370 | atg<br>Met        | cag<br>Gln        | agg<br>Arg        | ggc<br>Gly        | aac<br>Asn<br>375  | ttc<br>Phe        | agg<br>Arg        | aac<br>Asn        | cag<br>Gln             | agg<br>Arg<br>380 | aag<br>Lys        | aca<br>Thr        | gtg<br>Val        | aag<br>Lys        | 1152 |
| tgc<br>Cys<br>385 | ttc<br>Phe        | aac<br>Asn        | tgt<br>Cys        | Gly               | aag<br>Lys<br>390 | gtg<br>Val         | ggc<br>Gly        | cac<br>His        | att<br>Ile        | gcc<br>Ala<br>395      | aag<br>Lys        | aac<br>Asn        | tgt<br>Cys        | agg<br>Arg        | gcc<br>Ala<br>400 | 1200 |
| ccc<br>Pro        | agg<br>Arg        | aag<br>Lys        | aag<br>Lys        | ggc<br>Gly<br>405 | tgc<br>Cys        | tgg<br>Trp         | aag<br>Lys        | tgt<br>Cys        | ggc<br>Gly<br>410 | Lys                    | gag<br>Glu        | ggc<br>Gly        | cac<br>His        | cag<br>Gln<br>415 | atg<br>Met        | 1248 |
| aag<br>Lys        | gac<br>Asp        | tgc<br>Cys        | aat<br>Asn<br>420 | gag<br>Glu        | agg<br>Arg        | cag<br>Gln         | gcc<br>Ala        | aac<br>Asn<br>425 | ttc<br>Phe        | ctg<br>Leu             | ggc               | aaa<br>Lys        | atc<br>Ile<br>430 | tgg<br>Trp        | ccc<br>Pro        | 1296 |
| tcc<br>Ser        | cac<br>His        | aag<br>Lys<br>435 | Gly               | agg<br>Arg        | cct<br>Pro        | Gly                | aac<br>Asn<br>440 | ttc<br>Phe        | ctc<br>Leu        | cag<br>Gln             | tcc<br>Ser        | agg<br>Arg<br>445 | cct<br>Pro        | gag<br>Glu        | ccc<br>Pro        | 1344 |
| aca<br>Thr        | gcc<br>Ala<br>450 | Pro               | ccc<br>Pro        | gag<br>Glu        | gag<br>Glu        | tcc<br>Ser<br>455  | ttc<br>Phe        | agg<br>Arg        | ttť<br>Phe        | GJÀ<br>âââ             | gag<br>Glu<br>460 | gag<br>Glu        | aag<br>Lys        | acc<br>Thr        | acc<br>Thr        | 1392 |
| ccc<br>Pro<br>465 | agc<br>Ser        | cag<br>Gln        | aag<br>Lys        | cag<br>Gln        | gag<br>Glu<br>470 | Pro                | att<br>Ile        | gac<br>Asp        | aag<br>Lys        | gag<br>Glu<br>475      | ren               | tac<br>Tyr        | ccc<br>Pro        | ctg<br>Leu        | gcc<br>Ala<br>480 | 1440 |
| tcc<br>Ser        | ctg<br>Leu        | agg<br>Arg        | tcc<br>Ser        | ctg<br>Leu<br>485 | Phe               | ggc<br>Gly         | aac<br>Asn        | gac<br>Asp        | Pro               | Ser                    | tcc<br>Ser        | cag<br>Gln        | taa<br>*          | (SI               | D NO:36) D NO:37) | 1482 |

Figure 30 B

Figure 31

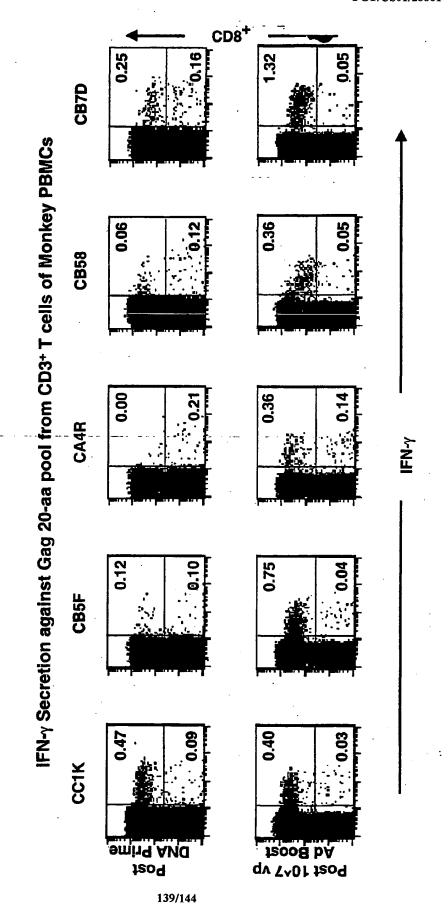
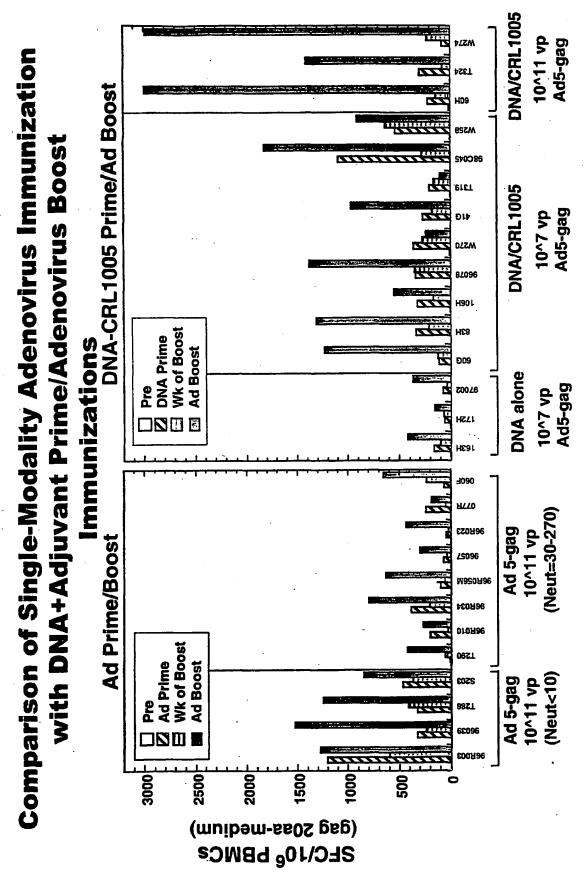


FIGURE 32



### FIGURE 33A

| ATGGGTGCTA | GGGCTTCTGT         | GCTGTCTGGT         | GGTGAGCTGG         | ACAAGTGGGA         | GAAGATCAGG |
|------------|--------------------|--------------------|--------------------|--------------------|------------|
| CTGAGGCCTG | GTGGCAAGAA         | GAAGTACAAG         | CTAAAGCACA         | TTGTGTGGGC         | CTCCAGGGAG |
| CTGGAGAGGT | TTGCTGTGAA         | CCCTGGCCTG         | CTGGAGACCT         | CTGAGGGGTG         | CAGGCAGATC |
| CTGGGCCAGC | TCCAGCCCTC         | CCTGCAAACA         | GGCTCTGAGG         | AGCTGAGGTC         | CCTGTACAAC |
| ACAGTGGCTA | CCCTGTACTG         | ${\tt TGTGCACCAG}$ | AAGATTGATG         | TGAAGGACAC         | CAAGGAGGCC |
| CTGGAGAAGA | TTGAGGAGGA         | GCAGAACAAG         | TCCAAGAAGA         | AGGCCCAGCA         | GGCTGCTGCT |
| GGCACAGGCA | ACTCCAGCCA         | GGTGTCCCAG         | AACTACCCCA         | TTGTGCAGAA         | CCTCCAGGGC |
| CAGATGGTGC | ACCAGGCCAT         | CTCCCCCGG          | ACCCTGAATG         | CCTGGGTGAA         | GGTGGTGGAG |
| GAGAAGGCCT | TCTCCCCTGA         | GGTGATCCCC         | ATGTTCTCTG         | CCCTGTCTGA         | GGGTGCCACC |
| CCCCAGGACC | TGAACACCAT         | GCTGAACACA         | GTGGGGGGCC         | ATCAGGCTGC         | CATGCAGATG |
| CTGAAGGAGA | CCATCAATGA         | GGAGGCTGCT         | GAGTGGGACA         | GGCTGCATCC         | TGTGCACGCT |
| GGCCCCATTG | CCCCGGCCA          | GATGAGGGAG         | CCCAGGGGCT         | ${\tt CTGACATTGC}$ | TGGCACCACC |
| TCCACCCTCC | AGGAGCAGAT         | ${\tt TGGCTGGATG}$ | ACCAACAACC         | CCCCCATCCC         | TGTGGGGGAA |
| ATCTACAAGA | ${\tt GGTGGATCAT}$ | CCTGGGCCTG         | AACAAGATTG         | ${\tt TGAGGATGTA}$ | CTCCCCCACC |
| TCCATCCTGG | ACATCAGGCA         | GGGCCCCAAG         | GAGCCCTTCA         | ${\tt GGGACTATGT}$ | GGACAGGTTC |
| TACAAGACCC | ${\tt TGAGGGCTGA}$ | ${\tt GCAGGCCTCC}$ | CAGGAGGTGA         | AGAACTGGAT         | GACAGAGACC |
| CTGCTGGTGC | AGAATGCCAA         | ${\tt CCCTGACTGC}$ | AAGACCATCC         | TGAAGGCCCT         | GGGCCCTGCT |
| GCCACCCTGG | AGGAGATGAT         | GACAGCCTGC         | CAGGGGGTGG         | GGGGCCCTGG         | TCACAAGGCC |
| AGGGTGCTGG | CTGAGGCCAT         | GTCCCAGGTG         | ACCAACTCCG         | CCACCATCAT         | GATGCAGAGG |
| GGCAACTTCA | GGAACCAGAG         | GAAGACAGTG         | AAGTGCTTCA         | ACTGTGGCAA         | GGTGGGCCAC |
| ATTGCCAAGA | ACTGTAGGGC         | CCCCAGGAAG         | AAGGGCTGCT         | ${\tt GGAAGTGTGG}$ | CAAGGAGGGC |
| CACCAGATGA | AGGACTGCAA         | TGAGAGGCAG         | ${\tt GCCAACTTCC}$ | TGGGCAAAAT         | CTGGCCCTCC |
| CACAAGGGCA | GGCCTGGCAA         | ${\tt CTTCCTCCAG}$ | TCCAGGCCTG         | AGCCCACAGC         | CCCTCCCGAG |
| GAGTCCTTCA | ${\tt GGTTTGGGGA}$ | GGAGAAGACC         | ACCCCCAGCC         | AGAAGCAGGA         | GCCCATTGAC |
| AAGGAGCTGT | ACCCCCTGGC         | ${\tt CTCCCTGAGG}$ | TCCCTGTTTG         | GCAACGACCC         | CTCCTCCCAG |
| ATGGCTCCCA | TCTCCCCCAT         | TGAGACTGTG         | ${\tt CCTGTGAAGC}$ | ${\tt TGAAGCCTGG}$ | CATGGATGGC |
| CCCAAGGTGA | AGCAGTGGCC         | CCTGACTGAG         | GAGAAGATCA         | AGGCCCTGGT         | GGAAATCTGC |
| ACTGAGATGG | AGAAGGAGGG         | CAAAATCTCC         | AAGATTGGCC         | CCGAGAACCC         | CTACAACACC |
| CCTGTGTTTG | CCATCAAGAA         | GAAGGACTCC         | ACCAAGTGGA         | GGAAGCTGGT         | GGACTTCAGG |
| GAGCTGAACA | AGAGGACCCA         | GGACTTCTGG         | GAGGTGCAGC         | TGGGCATCCC         | CCACCCCGCT |
| GGCCTGAAGA | AGAAGAAGTC         | TGTGACTGTG         | CTGGCTGTGG         | GGGATGCCTA         | CTTCTCTGTG |
| CCCCTGGATG | AGGACTTCAG         | GAAGTACACT         | GCCTTCACCA         | TCCCCTCCAT         | CAACAATGAG |
| ACCCCTGGCA | TCAGGTACCA         | GTACAATGTG         | CTGCCCCAGG         | GCTGGAAGGG         | CTCCCCTGCC |
| ATCTTCCAGT | CCTCCATGAC         | CAAGATCCTG         | GAGCCCTTCA         | GGAAGCAGAA         | CCCTGACATT |
| GTGATCTACC | AGTACATGGC         | TGCCCTGTAT         | GTGGGCTCTG         | ACCTGGAGAT         | TGGGCAGCAC |
| AGGACCAAGA | TTGAGGAGCT         | GAGGCAGCAC         | CTGCTGAGGT         | GGGGCCTGAC         | CACCCCTGAC |
| AAGAAGCACC | AGAAGGAGCC         | CCCCTTCCTG         | TGGATGGGCT         | ATGAGCTGCA         | CCCCGACAAG |
| TGGACTGTGC | AGCCCATTGT         | GCTGCCTGAG         | AAGGACTCCT         | GGACTGTGAA         | TGACATCCAG |
| AAGCTGGTGG | GCAAGCTGAA         | CTGGGCCTCC         | CAAATCTACC         | CTGGCATCAA         | GGTGAGGCAG |
| CTGTGCAAGC | TGCTGAGGGG         | CACCAAGGCC         | CTGACTGAGG         | TGATCCCCCT         | GACTGAGGAG |
| GCTGAGCTGG | AGCTGGCTGA         | GAACAGGGAG         | ATCCTGAAGG         | AGCCTGTGCA         | TGGGGTGTAC |

# FIGURE 33B

| TATGACCCCT | CCAAGGACCT | GATTGCTGAG | ATCCAGAAGC         | AGGGCCAGGG | CCAGTGGACC |
|------------|------------|------------|--------------------|------------|------------|
| TACCAAATCT | ACCAGGAGCC | CTTCAAGAAC | ${\tt CTGAAGACTG}$ | GCAAGTATGC | CAGGATGAGG |
| GGGCCCACA  | CCAATGATGT | GAAGCAGCTG | ACTGAGGCTG         | TGCAGAAGAT | CACCACTGAG |
| TCCATTGTGA | TCTGGGGCAA | GACCCCCAAG | ${\tt TTCAAGCTGC}$ | CCATCCAGAA | GGAGACCTGG |
| GAGACCTGGT | GGACTGAGTA | CTGGCAGGCC | ACCTGGATCC         | CTGAGTGGGA | GTTTGTGAAC |
| ACCCCCCCC  | TGGTGAAGCT | GTGGTACCAG | CTGGAGAAGG         | AGCCCATTGT | GGGGGCTGAG |
| ACCTTCTATG | TGGCTGGGGC | TGCCAACAGG | GAGACCAAGC         | TGGGCAAGGC | TGGCTATGTG |
| ACCAACAGGG | GCAGGCAGAA | GGTGGTGACC | CTGACTGACA         | CCACCAACCA | GAAGACTGCC |
| CTCCAGGCCA | TCTACCTGGC | CCTCCAGGAC | TCTGGCCTGG         | AGGTGAACAT | TGTGACTGCC |
| TCCCAGTATG | CCCTGGGCAT | CATCCAGGCC | CAGCCTGATC         | AGTCTGAGTC | TGAGCTGGTG |
| AACCAGATCA | TTGAGCAGCT | GATCAAGAAG | GAGAAGGTGT         | ACCTGGCCTG | GGTGCCTGCC |
| CACAAGGGCA | TTGGGGGCAA | TGAGCAGGTG | GACAAGCTGG         | TGTCTGCTGG | CATCAGGAAG |
| GTGCTGTTCC | TGGATGGCAT | TGACAAGGCC | CAGGATGAGC         | ATGAGAAGTA | CCACTCCAAC |
| TGGAGGGCTA | TGGCCTCTGA | CTTCAACCTG | CCCCTGTGG          | TGGCTAAGGA | GATTGTGGCC |
| TCCTGTGACA | AGTGCCAGCT | GAAGGGGGAG | GCCATGCATG         | GGCAGGTGGA | CTGCTCCCCT |
| GGCATCTGGC | AGCTGGCCTG | CACCCACCTG | GAGGGCAAGG         | TGATCCTGGT | GGCTGTGCAT |
| GTGGCCTCCG | GCTACATTGA | GGCTGAGGTG | ATCCCTGCTG         | AGACAGGCCA | GGAGACTGCC |
| TACTTCCTGC | TGAAGCTGGC | TGGCAGGTGG | CCTGTGAAGA         | CCATCCACAC | TGCCAATGGC |
| TCCAACTTCA | CTGGGGCCAC | AGTGAGGGCT | GCCTGCTGGT         | GGGCTGGCAT | CAAGCAGGAG |
| TTTGGCATCC | CCTACAACCC | CCAGTCCCAG | GGGGTGGTGG         | CCTCCATGAA | CAAGGAGCTG |
| AAGAAGATCA | TTGGGCAGGT | GAGGGACCAG | GCTGAGCACC         | TGAAGACAGC | TGTGCAGATG |
| GCTGTGTTCA | TCCACAACTT | CAAGAGGAAG | GGGGGCATCG         | GGGGCTACTC | CGCTGGGGAG |
| AGGATTGTGG | ACATCATTGC | CACAGACATC | CAGACCAAGG         | AGCTCCAGAA | GCAGATCACC |
| AAGATCCAGA | ACTTCAGGGT | GTACTACAGG | GACTCCAGGA         | ACCCCCTGTG | GAAGGGCCCT |
| GCCAAGCTGC | TGTGGAAGGG | GGAGGGGGCT | GTGGTGATCC         | AGGACAACTC | TGACATCAAG |
| GTGGTGCCCA | GGAGGAAGGC | CAAGATCATC | AGGGACTATG         | GCAAGCAGAT | GGCTGGGGAT |
| GACTGTGTGG | CCTCCAGGCA | GGATGAGGAC | TAA .              |            |            |
| SEQ ID NO: | 38         |            |                    |            |            |

#### FIGURE 34A

Met Gly Ala Arg Ala Ser Val Leu Ser Gly Gly Glu Leu Asp Lys Trp Glu Lys Ile Arg Leu Arg Pro Gly Gly Lys Lys Lys Tyr Lys Leu Lys His Ile Val Trp Ala Ser Arg Glu Leu Glu Arg Phe Ala Val Asn Pro Gly Leu Leu Glu Thr Ser Glu Gly Cys Arg Gln Ile Leu Gly Gln Leu Gln Pro Ser Leu Gln Thr Gly Ser Glu Glu Leu Arg Ser Leu Tyr Asn Thr Val Ala Thr Leu Tyr Cys Val His Gln Lys Ile Asp Val Lys Asp Thr Lys Glu Ala Leu Glu Lys Ile Glu Glu Glu Gln Asn Lys Ser Lys Lys Lys Ala Gln Gln Ala Ala Gly Thr Gly Asn Ser Ser Gln Val Ser Gln Asn Tyr Pro Ile Val Gln Asn Leu Gln Gly Gln Met Val His Gln Ala Ile Ser Pro Arg Thr Leu Asn Ala Trp Val Lys Val Val Glu Glu Lys Ala Phe Ser Pro Glu Val Ile Pro Met Phe Ser Ala Leu Ser Glu Gly Ala Thr Pro Gln Asp Leu Asn Thr Met Leu Asn Thr Val Gly Gly His Gln Ala Ala Met Gln Met Leu Lys Glu Thr Ile Asn Glu Glu Ala Ala Glu Trp Asp Arg Leu His Pro Val His Ala Gly Pro Ile Ala Pro Gly Gln Met Arg Glu Pro Arg Gly Ser Asp Ile Ala Gly Thr Thr Ser Thr Leu Gln Glu Gln Ile Gly Trp Met Thr Asn Asn Pro Pro Ile Pro Val Gly Glu Ile Tyr Lys Arg Trp Ile Ile Leu Gly Leu Asn Lys Ile Val Arg Met Tyr Ser Pro Thr Ser Ile Leu Asp Ile Arg Gln Gly Pro Lys Glu Pro Phe Arg Asp Tyr Val Asp Arg Phe Tyr Lys Thr Leu Arg Ala Glu Gln Ala Ser Gln Glu Val Lys Asn Trp Met Thr Glu Thr Leu Leu Val Gln Asn Ala Asn Pro Asp Cys Lys Thr Ile Leu Lys Ala Leu Gly Pro Ala Ala Thr Leu Glu Glu Met Met Thr Ala Cys Gln Gly Val Gly Gly Pro Gly His Lys Ala Arg Val Leu Ala Glu Ala Met Ser Gln Val Thr Asn Ser Ala Thr Ile Met Met Gln Arg Gly Asn Phe Arg Asn Gln Arg Lys Thr Val Lys Cys Phe Asn Cys Gly Lys Val Gly His Ile Ala Lys Asn Cys Arg Ala Pro Arg Lys Lys Gly Cys Trp Lys Cys Gly Lys Glu Gly His Gln Met Lys Asp Cys Asn Glu Arg Gln Ala Asn Phe Leu Gly Lys Ile Trp Pro Ser His Lys Gly Arg Pro Gly Asn Phe Leu Gln Ser Arg Pro Glu Pro Thr Ala Pro Pro Glu Glu Ser Phe Arg Phe Gly Glu Glu Lys Thr Thr Pro Ser Gln Lys Gln Glu Pro Ile Asp Lys Glu Leu Tyr Pro Leu Ala Ser Leu Arg Ser Leu Phe Gly Asn Asp Pro Ser Ser Gln Met Ala Pro Ile Ser Pro Ile Glu Thr Val Pro Val Lys Leu Lys Pro Gly Met Asp Gly Pro Lys Val Lys Gln Trp Pro Leu Thr Glu Glu Lys Ile Lys Ala Leu Val Glu Ile Cys Thr Glu Met Glu Lys Glu Gly Lys Ile Ser Lys Ile Gly Pro Glu Asn Pro Tyr Asn Thr Pro Val Phe Ala Ile Lys Lys Asp Ser Thr Lys Trp Arg Lys Leu Val Asp Phe Arg Glu Leu Asn Lys Arg Thr Gln Asp Phe Trp Glu Val Gln Leu Gly Ile Pro His Pro Ala Gly Leu Lys Lys Lys Ser Val Thr Val Leu Ala Val Gly Asp Ala Tyr Phe Ser Val Pro Leu Asp Glu Asp Phe Arg Lys Tyr Thr Ala Phe Thr Ile Pro Ser Ile Asn Asn Glu Thr Pro Gly Ile Arg Tyr Gln Tyr Asn Val Leu Pro Gln Gly Trp Lys Gly Ser Pro Ala Ile Phe Gln Ser Ser Met Thr Lys Ile Leu Glu Pro Phe Arg Lys Gln Asn Pro Asp Ile Val Ile Tyr Gln Tyr Met Ala Ala Leu Tyr Val Gly Ser Asp Leu Glu Ile Gly Gln His Arg Thr Lys Ile Glu Glu Leu Arg Gln His Leu Leu Arg Trp Gly Leu Thr Thr Pro Asp Lys Lys His Gln Lys Glu Pro Pro Phe Leu Trp Met Gly Tyr Glu Leu His Pro

#### FIGURE 34B

Asp Lys Trp Thr Val Gln Pro Ile Val Leu Pro Glu Lys Asp Ser Trp Thr Val Asn Asp Ile Gln Lys Leu Val Gly Lys Leu Asn Trp Ala Ser Gln Ile Tyr Pro Gly Ile Lys Val Arg Gln Leu Cys Lys Leu Leu Arg Gly Thr Lys Ala Leu Thr Glu Val Ile Pro Leu Thr Glu Glu Ala Glu Leu Glu Leu Ala Glu Asn Arg Glu Ile Leu Lys Glu Pro Val His Gly Val Tyr Tyr Asp Pro Ser Lys Asp Leu Ile Ala Glu Ile Gln Lys Gln Gly Gln Gly Gln Trp Thr Tyr Gln Ile Tyr Gln Glu Pro Phe Lys Asn Leu Lys Thr Gly Lys Tyr Ala Arg Met Arg Gly Ala His Thr Asn Asp Val Lys Gln Leu Thr Glu Ala Val Gln Lys Ile Thr Thr Glu Ser Ile Val Ile Trp Gly Lys Thr Pro Lys Phe Lys Leu Pro Ile Gln Lys Glu Thr Trp Glu Thr Trp Trp Thr Glu Tyr Trp Gln Ala Thr Trp Ile Pro Glu Trp Glu Phe Val Asn Thr Pro Pro Leu Val Lys Leu Trp Tyr Gln Leu Glu Lys Glu Pro Ile Val Gly Ala Glu Thr Phe Tyr Val Ala Gly Ala Ala Asn Arg Glu Thr Lys Leu Gly Lys Ala Gly Tyr Val Thr Asn Arg Gly Arg Gln Lys Val Val Thr Leu Thr Asp Thr Thr Asn Gln Lys Thr Ala Leu Gln Ala Ile Tyr Leu Ala Leu Gln Asp Ser Gly Leu Glu Val Asn Ile Val Thr Ala Ser Gln Tyr Ala Leu Gly Ile Ile Gln Ala Gln Pro Asp Gln Ser Glu Ser Glu Leu Val Asn Gln Ile Ile Glu Gln Leu Ile Lys Lys Glu Lys Val Tyr Leu Ala Trp Val Pro Ala His Lys Gly Ile Cly Gly Asn Glu Gln Val Asp Lys Leu Val Ser Ala Gly Ile Arg Lys Val Leu Phe Leu Asp Gly Ile Asp Lys Ala Gln Asp Glu His Glu Lys Tyr His Ser Asn Trp Arg Ala Met Ala Ser Asp Phe Asn Leu Pro Pro Val Val Ala Lys Glu Ile Val Ala Ser Cys Asp Lys Cys Gln Leu Lys Gly Glu Ala Met His Gly Gln Val Asp Cys Ser Pro Gly Ile Trp Gln Leu Ala Cys Thr His Leu Glu Gly Lys Val Ile Leu Val Ala Val His Val Ala Ser Gly Tyr Ile Glu Ala Glu Val Ile Pro Ala Glu Thr Gly Gln Glu Thr Ala Tyr Phe Leu Leu Lys Leu Ala Gly Arg Trp Pro Val Lys Thr Ile His Thr Ala Asn Gly Ser Asn Phe Thr Gly Ala Thr Val Arg Ala Ala Cys Trp Trp Ala Gly Ile Lys Gln Glu Phe Gly Ile Pro Tyr Asn Pro Gln Ser Gln Gly Val Val Ala Ser Met Asn Lys Glu Leu Lys Lys Ile Ile Gly Gln Val Arg Asp Gln Ala Glu His Leu Lys Thr Ala Val Gln Met Ala Val Phe Ile His Asn Phe Lys Arg Lys Gly Gly Jle Gly Gly Tyr Ser Ala Gly Glu Arg Ile Val Asp Ile Ile Ala Thr Asp Ile Gln Thr Lys Glu Leu Gln Lys Gln Ile Thr Lys Ile Gln Asn Phe Arg Val Tyr Tyr Arg Asp Ser Arg Asn Pro Leu Trp Lys Gly Pro Ala Lys Leu Leu Trp Lys Gly Glu Gly Ala Val Val Ile Gln Asp Asn Ser Asp Ile Lys Val Val Pro Arg Arg Lys Ala Lys Ile Ile Arg Asp Tyr Gly Lys Gln Met Ala Gly Asp Asp Cys Val Ala Ser Arg Gln Asp Glu Asp SEQ ID NO: 39

#### (19) World Intellectual Property Organization International Bureau



# 

#### (43) International Publication Date 21 March 2002 (21.03.2002)

#### **PCT**

### (10) International Publication Number WO 02/22080 A3

(51) International Patent Classification7:

- C12N 15/86
- (21) International Application Number: PCT/US01/28861
- (22) International Filing Date:

14 September 2001 (14.09.2001)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

60/233,180

15 September 2000 (15.09.2000)

- (71) Applicant (for all designated States except US): MERCK & CO., INC. [US/US]: 126 East Lincoln Avenue, Rahway, NJ 07065-0907 (US).
- (72) Inventors; and
- (75) Inventors/Applicants (for US only): EMINI, Emilio, A. [US/US]: 126 East Lincoln Avenue, Rahway, NJ 07065-0907 (US). YOUIL, Rima [AU/US]; 126 East Lincoln Avenue, Rahway, NJ 07065-0907 (US). BETT, Andrew, J. [CA/US]; 126 East Lincoln Avenue, Rahway, NJ 07065-0907 (US). CHEN, Ling [US/US]: 126 East Lincoln Avenue. Rahway. NJ 07065-0907 (US). KASLOW, David, C. [US/US]: 126 East Lincoln Avenue, Rahway, NJ 07065-. 0907 (US). SHIVER, John, W. [US/US]; 126 East Lincoln Avenue. Rahway, NJ 07065-0907 (US). TONER, Timothy, J. [US/US]: 126 East Lincoln Avenue, Rahway, NJ 07065-0907 (US). CASIMIRO, Daniel, R. [PH/US]: 126 East Lincoln Avenue, Rahway, NJ 07065-0907 (US).

- (74) Common Representative: MERCK & CO., INC.; 126 East Lincoln Avenue, Rahway, NJ 07065-0907 (US).
- (81) Designated States (national): AE. AG. AL. AM. AT. AU. AZ. BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ. DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM. HR. HU, ID. IL. IN. IS. JP. KE. KG. KR, KZ. LC, LK. LR. LS. LT. LU. LV. MA. MD. MG, MK. MN. MW, MX, MZ, NO, NZ, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL. TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.
- (84) Designated States (regional): ARIPO patent (GH, GM, KE. LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT. LU. MC. NL. PT. SE, TR), OAPI patent (BF, BJ, CF, CG. CI. CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD. TG).

#### Published:

- with international search report
- before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments
- (88) Date of publication of the international search report: 2 May 2002

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: ENHANCED FIRST GENERATION ADENOVIRUS VACCINES EXPRESSING CODON OPTIMIZED HIVI-GAG. POL. NEF AND MODIFICATIONS

(57) Abstract: First generation adenoviral vectors and associated recombinant adenovirus-based HIV vaccines which show enhanced stability and growth properties and greater cellular-mediated immunity are described within this specification. These adenoviral vectors are utilized to generate and produce through cell culture various adenoviral-based HIV-1 vaccines which contain HIV-I gag, HIV-I pol and/or HIV-I nef polynucleotide pharmaceutical products, and biologically relevant modifications thereof. These adenovirus vaccines, when directly introduced into living vertebrate tissue, preferably a mammalian host such as a human or a non-human mammal of commercial or domestic veterinary importance, express the HIV1- Gag. Pol and/or Ncf protein or biologically modification thereof, inducing a cellular immune response which specifically recognizes HIV-1. The exemplified polynucleotides of the present invention are synthetic DNA molecules encoding HIV-1 Gag, encoding codon optimized HIV-1 Pol, derivatives of optimized HIV-1 Pol (including constructs wherein protease, reverse transcriptase, RNAse H and integrase activity of HIV-1 Pol is inactivated). HIV-1 Nef and derivatives of optimized HIV-1 Nef, including nef mutants which effect wild type characteristics of Nef. such as myristylation and down regulation of host CD4. The adenoviral vaccines of the present invention, when administered alone or in a combined modality regime, will offer a prophylactic advantage to previously uninfected individuals and/or provide a therapeutic effect by reducing viral load levels within an infected individual, thus prolonging the asymptomatic phase of HIV-I infection.

International application No.

PCT/US01/28861

| IPC(7) US CL According to B. FIELI Minimum doc U.S.: 42 | SIFICATION OF SUBJECT MATTER  : C12N 15/86  : 435/456  International Patent Classification (IPC) or to both na OS SEARCHED  numentation searched (classification system followed b 4/205.1, 207.1, 227.1, 233.1; 435/69.1, 69.3, 173.3, on searched other than minimum documentation to the | y classification symbols)<br>, 235.1, 320.1, 456; 530/23.72;  | in the fields searched  |  |  |
|---|---|---|---|--|--|
| Please See Co   | ta base consulted during the international search (name   | e of data base and, where practicable, s  | earch terms used)   |  |  |
| C. DOCT   | JMENTS CONSIDERED TO BE RELEVANT  |   |   |  |  |
| Category *  | Citation of document, with indication, where app  | propriate, of the relevant passages   | Relevant to claim No.   |  |  |
| X<br><br>Y  | WO 96/39178 (ERTL et al.) 12 December 1996 (12. and claims 1 and 5.   | 12.1996), see page 5, 6,10, 12, 13  | 1-3, 8-11, 18<br><br>4, 5, 13-17, 29, 30,<br>32, 34, 35, 37                     |  |  |
| Х<br>—<br>Y   | US 6,019,978 A (ERTL et al.) 1 February 2000 (01/   | (02/2000), see columns 2, 7 and 8.  | 1-3, 8-11, 18<br><br>4, 5, 13-17, 29, 30, 32, 34, 35, 37                        |  |  |
| X,P<br>X<br>  | US 6,287,571 A A (ERTL et al.) 11 September 2000<br>and claim 1.<br>US 5,643,579A (HUNG et al.) 1 July 1997 (01/07/1  | •   | 1, 9, 18<br>1-3, 8, 9-11, 18<br>  |  |  |
| Y   | WANG et al. The use of an E1-deleted, replication expressing the rabies virus glycoprotein for early va Journal of Virology (March 1997) Vol. 71, No. 5, p  | ectination of mice against rables virus.  | 34, 35, 37<br>1-3, 9-11, 13-18  |  |  |
| Furthe  | or documents are listed in the continuation of Box C.   | See patent family annex.  |   |  |  |
| "A" documer of partic                                   | Special categories of cited documents:  It defining the general state of the art which is not considered to be cular relevance  Application or patent published on or after the international filing date   | "T" later document published after the int date and not in conflict with the appli principle or theory underlying the int document of particular relevance; the considered novel or cannot be considered novel or taken alone | cation but cited to understand the vention cention celaimed invention cannot be |  |  |
| establish<br>specifie                                   | establish the publication date of another citation or other special reason (8 considered to involve an inventive step when the document is specified)  combined with one or more other such documents, such combination being obvious to a person skilled in the art                        |   |   |  |  |
| *P* docume<br>priority                                  | "O" document reterring to an oral materials. The comment member of the same parent family priority date claimed   |   |   |  |  |
|   | actual completion of the international search   | 13 MAR 2002   |   |  |  |
| Name and  | y 2002 (06.02.2002) mailing address of the ISA/US commissioner of Patents and Trademarks  | Authorized officer  | 1/0   |  |  |
| B. W.   | Box PCT Washington, D.C. 20231 Talanhara No. 703 208 0196   |   |   |  |  |
| 1   | No. (703)305-3230   | <u> </u>  |   |  |  |

International application No.

PCT/US01/28861

# INTERNATIONAL SEARCH REPORT

| ategory *  | Citation of document, with indication, where appropriate, of the relevant passages  | Relevant to claim No |
|------------|---|----------------------|
| Y          | NATUK et al. Immunogenicity of recombinant human adenovirus -human immunodeficiency virus vaccines in chimpanzees. Aids Research and Human Retroviruses (1993) Vol. 9, No. 5, pp395-404, see material and methods.  | 1, 9, 29, 30, 32     |
| Y          | PREVEC et al. Immune response to HIV-1 gag antigens induced by recombinant adenovirus vectors in mice and rhesus macaque monkeys. Journal of Acquired Immune Deficincy Syndrome. (1991) Vol. 4, No. 6 pp. 568-76, see abstract.                           | 1, 9, 29, 30, 32     |
| <b>, Y</b> | LORI et al. Rapid protection against human immunodeficiency virus type 1 (HIV-1) replication mediated by high efficiency non-retroviral delivery of genes interfering with HIV-1 tat and gag. Gene Therapy (1994) Vol. 1, No. 1, pp. 27-31, see abstract. | 1, 9                 |
| Y          | PFARR et al. Differential effects of polyadenylation regions on gene expression in mammalian cells. DNA (1986) Vol. 5, No. 2, pp.115-22, see abstract.  | . 16                 |
| Y          | NATUK et al. Adenovirus vectored vaccine. Developmental Biological Standards (1994)<br>Vol. 82, pp. 71-77, see abstract.  | 1, 9                 |
|            |   |                      |
|            |   |                      |
|            | · · · · · · · · · · · · · · · · · · ·   |                      |
|            |   |                      |
|            |   |                      |
|            |   |                      |
|            |   |                      |
|            |   |                      |
|            |   |                      |
|            |   |                      |
|            |   |                      |
|            |   |                      |
|            |   |                      |
|            |   | ·                    |
|            |   |                      |

International application No.

PCT/US01/28861

| Box  | Box I Observations where certain claims were found unsearchable (Continuation of Item 1 of first sheet)                           |  |  |  |  |
|------|---|--|--|--|--|
| This | This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons: |  |  |  |  |
| 1.   |   | Claim Nos.: because they relate to subject matter not required to be searched by this Authority, namely:   |  |  |  |
| 2.   |   | Claim Nos.: 31 because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically: This claim could not be searched because applicant did not provide a CRF. |  |  |  |
|      | _   |  |  |  |  |
| 3.   |   | Claim Nos: because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule   |  |  |  |
|      | 6.4(a).   |  |  |  |  |
| Box  | II Ob   | servations where unity of invention is lacking (Continuation of Item 2 of first sheet)   |  |  |  |
| This | Internat  | ional Searching Authority found multiple inventions in this international application, as follows:   |  |  |  |
| Plea | ise See C   | ontimuation Sheet  |  |  |  |
|      |   |  |  |  |  |
|      |   |  |  |  |  |
|      |   |  |  |  |  |
|      |   | ·  |  |  |  |
| 1.   |   | As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.   |  |  |  |
| 2.   |   | As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.   |  |  |  |
| 3.   |   | As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:   |  |  |  |
|      |   | report covers only diose crams for which rees were place, specifically crams from  |  |  |  |
|      |   |  |  |  |  |
|      |   |  |  |  |  |
|      |   |  |  |  |  |
| 4.   | $\boxtimes$   | No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.: 1-5, 8-11, 13-18, 29-32, 34, 35, 37   |  |  |  |
| Re   | mark on   | Protest The additional search fees were accompanied by the applicant's protest.  |  |  |  |
|      |   | No protest accompanied the payment of additional search fees.  |  |  |  |
| 1    |   |  |  |  |  |

Form PCT/ISA/210 (continuation of first sheet(1)) (July 1998)

International application No.

PCT/US01/28861

The special technical feature of group 4, 16 and 31 is considered to be a method of producing recombinant adenoviral particles. Each group contains different sequences hence the resulting particles would have different structures and functions associated with the particle.

The special technical feature of group 5, 17 and 32 is considered to be a method of producing a cellular mediated immune response to the heterologous protein encoded by the different adenoviral vectors. Each group contains different sequences a encoding different protein, therefore the resulting immune response will also be different.

The special technical feature of group 6, 18 and 33 is considered to be a method of producing a cellular mediated immune response to the heterologous protein encoded by the different adenoviral vectors in conjunction with immunizing the individual a DNA plasmid vaccine. Each method contains different sequences encoding a different protein, therefore the resulting immune response will also be different.

Accordingly, groups 1-48 are not so linked by the same or corresponding technical feature as to form a single general inventive concept.

Continuation of B. FIELDS SEARCHED Item 3:

WEST 2.0, STN-BIOSIS, MEDLINE

adenoviral vector, deletion, HIV, Gag, polyadenylation signal, CMV promoter

International application No.

PCT/US01/28861

|        |            | and ΔE3, the vector contains the cis-acting packaging sequence of the wild type adenovirus genome, and a gene which encodes an HIV Pol protein (SEQ ID NO: 1)                            |
|--------|------------|--|
| _      |            | inserted in E1.  |
| 4      | 55         | The claim is directed to an adenoviral vector that is at least partially deleted of $\Delta E1$ and $\Delta E3$ , the vector contains the cis-acting packaging sequence of the wild type |
|        |            | adenovirus genome, and a gene which encodes an HIV Pol protein (SEQ ID NO: 5) inserted in E1.  |
| 5      | 55         | The claim is directed to an adenoviral vector that is at least partially deleted of $\Delta E1$  |
| ,      |            | and $\Delta E3$ , the vector contains the cis-acting packaging sequence of the wild type adenovirus genome, and a gene which encodes an HIV Pol protein (SEQ ID NO: 7) inserted in E1.   |
| 6      | 57-61      | The claims are directed to a method of making and harvesting of a recombinant  |
| 7      | 62, 65, 66 | adenoviral particle that contains a gene encoding an HIV Pol protein.  The claim is directed to a method of generating a cellular mediated immune response                               |
|        | 02, 03, 00 | to HIV Pol protein with the recombinant adenoviral particle.   |
| 18     | 63, 64     | The claim is directed to a method of generating a cellular mediated immune response  |
|        | 1          | to HIV Pol protein with the recombinant adenoviral particle in addition to administering a DNA plasmid vaccine.  |
| <br>19 | 67-70, 72, | The claims are directed to an adenoviral vector that is at least partially deleted of  |
|        | 73, 75     | ΔE1, the vector contains the cis-acting packaging sequence of the wild type  |
|        | '          | adenovirus genome, and a gene which encodes an HIV Nef protein (SEQ ID NO: 9)  |
|        |            | inserted in the parallel orientation of E1.  The claims are directed to an adenoviral vector that is at least partially deleted of   |
| 20     | 67-70, 72, | The claims are directed to an adenoviral vector that is at least partially defected of ΔΕ1, the vector contains the cis-acting packaging sequence of the wild type                       |
|        | 73, 75     | adenovirus genome, and a gene which encodes an HIV Nef protein (SEQ ID NO: 11)   |
|        |            | inserted in the parallel orientation of E1.  |
| 21     | 67-70, 72, | The claims are directed to an adenoviral vector that is at least partially deleted of  |
|        | 73, 75     | $\Delta E1$ , the vector contains the cis-acting packaging sequence of the wild type   |
|        |            | adenovirus genome, and a gene which encodes an HIV Nef protein (SEQ ID NO: 13) inserted in the parallel orientation of E1.   |
| 22     | 67-70, 72, | The claims are directed to an adenoviral vector that is at least partially deleted of  |
|        | 73, 75     | ΔE1, the vector contains the cis-acting packaging sequence of the wild type  |
|        | ,          | adenovirus genome, and a gene which encodes an HIV Nef protein (SEQ ID NO: 15)   |
|        |            | inserted in the parallel orientation of E1.  |
| 23     | 71         | The claim is directed to an adenoviral vector that is at least partially deleted of $\Delta E1$ , the vector contains the cis-acting packaging sequence of the wild type adenovirus      |
|        |            | genome, and a gene which encodes an HIV Nef protein (SEQ ID NO: 9) inserted in   |
|        |            | the antiparallel orientation of E1.  |
| 24     | 71         | The claim is directed to an adenoviral vector that is at least partially deleted of $\Delta E1$ ,  |
|        |            | the vector contains the cis-acting packaging sequence of the wild type adenovirus  |
|        |            | genome, and a gene which encodes an HIV Nef protein (SEQ ID NO: 11) inserted in  |
| 26     | 71         | the antiparallel orientation of E1.  The claim is directed to an adenoviral vector that is at least partially deleted of $\Delta E1$ ,   |
| 25     | /'         | the vector contains the cis-acting packaging sequence of the wild type adenovirus  |
|        |            | genome, and a gene which encodes an HIV Nef protein (SEQ ID NO: 13) inserted in  |
|        |            | the antiparallel orientation of E1.  |
| 26     | 71         | The claim is directed to an adenoviral vector that is at least partially deleted of $\Delta E1$ , the vector contains the cis-acting packaging sequence of the wild type adenovirus      |
|        |            | genome, and a gene which encodes an HIV Nef protein (SEQ ID NO: 15) inserted in  |
|        |            | the antiparallel orientation of E1.  |
| 27     | 74         | The claim is directed to an adenoviral vector that is at least partially deleted of $\Delta E1$  |
|        |            | and $\Delta E3$ , the vector contains the cis-acting packaging sequence of the wild type   |
|        | j          | adenovirus genome, and a gene which encodes an HIV Nef protein (SEQ ID NO: 9)  |
|        |            | inserted in E1.  |
| 28     | 74         | The claim is directed to an adenoviral vector that is at least partially deleted of $\Delta E1$ and $\Delta E3$ , the vector contains the cis-acting packaging sequence of the wild type |
|        |            | and AE3, the vector contains the cis-acting packaging sequence of the whotype adenovirus genome, and a gene which encodes an HIV Nef protein (SEQ ID NO: 11)                             |
|        | ļ          | inserted in E1.  |
| 29     | 74         | The claim is directed to an adenoviral vector that is at least partially deleted of $\Delta EI$  |
|        | 1          | and $\Delta E3$ , the vector contains the cis-acting packaging sequence of the wild type   |

International application No.

PCT/US01/28861

|    |                      | and $\Delta E3$ , the vector contains the cis-acting packaging sequence of the wild type adenovirus genome, and a gene which encodes an HIV Pol protein (SEQ ID NO: 1) inserted in E1.   |
|----|----------------------|--|
| 14 | 55                   | The claim is directed to an adenoviral vector that is at least partially deleted of $\Delta E1$ and $\Delta E3$ , the vector contains the cis-acting packaging sequence of the wild type adenovirus genome, and a gene which encodes an HIV Pol protein (SEQ ID NO: 5) inserted in E1.                         |
| 15 | 55                   | The claim is directed to an adenoviral vector that is at least partially deleted of $\Delta E1$ and $\Delta E3$ , the vector contains the cis-acting packaging sequence of the wild type adenovirus genome, and a gene which encodes an HIV Pol protein (SEQ ID NO: 7) inserted in E1.                         |
| 16 | 57-61                | The claims are directed to a method of making and harvesting of a recombinant adenoviral particle that contains a gene encoding an HIV Pol protein.  |
| 17 | 62, 65, 66           | The claim is directed to a method of generating a cellular mediated immune response to HIV Pol protein with the recombinant adenoviral particle.   |
| 18 | 63, 64               | The claim is directed to a method of generating a cellular mediated immune response to HIV Pol protein with the recombinant adenoviral particle in addition to administering a DNA plasmid vaccine.  |
| 19 | 67-70, 72,<br>73, 75 | The claims are directed to an adenoviral vector that is at least partially deleted of<br><u>AE1</u> , the vector contains the cis-acting packaging sequence of the wild type<br>adenovirus genome, and a gene which encodes an HIV Nef protein (SEQ ID NO: 9)<br>inserted in the parallel orientation of E1.   |
| 20 | 67-70, 72,<br>73, 75 | The claims are directed to an adenoviral vector that is at least partially deleted of<br><u>AE1</u> , the vector contains the cis-acting packaging sequence of the wild type adenovirus genome, and a gene which encodes an <u>HIV Nef protein (SEQ ID NO: 11)</u> inserted in the parallel orientation of E1. |
| 21 | 67-70, 72,<br>73, 75 | The claims are directed to an adenoviral vector that is at least partially deleted of<br><u>AE1</u> , the vector contains the cis-acting packaging sequence of the wild type adenovirus genome, and a gene which encodes an HIV Nef protein (SEQ ID NO: 13) inserted in the parallel orientation of E1.        |
| 22 | 67-70, 72,<br>73, 75 | The claims are directed to an adenoviral vector that is at least partially deleted of ΔΕ1, the vector contains the cis-acting packaging sequence of the wild type adenovirus genome, and a gene which encodes an HIV Nef protein (SEQ ID NO: 15) inserted in the parallel orientation of Ε1.                   |
| 23 | 71                   | The claim is directed to an adenoviral vector that is at least partially deleted of $\Delta E1$ , the vector contains the cis-acting packaging sequence of the wild type adenovirus genome, and a gene which encodes an HIV Nef protein (SEQ ID NO: 9) inserted in the antiparallel orientation of E1.         |
| 24 | 71                   | The claim is directed to an adenoviral vector that is at least partially deleted of $\Delta E1$ , the vector contains the cis-acting packaging sequence of the wild type adenovirus genome, and a gene which encodes an HIV Nef protein (SEQ ID NO: 11) inserted in the antiparallel orientation of E1.        |
| 25 | 71                   | The claim is directed to an adenoviral vector that is at least partially deleted of $\Delta E1$ , the vector contains the cis-acting packaging sequence of the wild type adenovirus genome, and a gene which encodes an HIV Nef protein (SEQ ID NO: 13) inserted in the antiparallel orientation of E1.        |
| 26 | 71                   | The claim is directed to an adenoviral vector that is at least partially deleted of $\Delta E1$ , the vector contains the cis-acting packaging sequence of the wild type adenovirus genome, and a gene which encodes an HIV Nef protein (SEQ ID NO: 15) inserted in the antiparallel orientation of E1.        |
| 27 | 74                   | The claim is directed to an adenoviral vector that is at least partially deleted of $\Delta E1$ and $\Delta E3$ , the vector contains the cis-acting packaging sequence of the wild type adenovirus genome, and a gene which encodes an HIV Nef protein (SEQ ID NO: 9) inserted in E1.                         |
| 28 | 74                   | The claim is directed to an adenoviral vector that is at least partially deleted of $\Delta E1$ and $\Delta E3$ , the vector contains the cis-acting packaging sequence of the wild type adenovirus genome, and a gene which encodes an HIV Nef protein (SEQ ID NO: 11) inserted in E1.                        |
|    | 74                   | The claim is directed to an adenoviral vector that is at least partially deleted of $\Delta E1$  |

International application No.

PCT/US01/28861

|                 | _           |  |
|-----------------|-------------|--|
|                 |             | adenovirus genome, and a gene which encodes an HIV Nef protein (SEQ ID NO: 13)   |
|                 | ļ           | invested in F1   |
|                 |             | The claim is directed to an adenoviral vector that is at least partially deleted of $\Delta E$ !   |
| 30              | 74          | the vector contains the cis-acting packaging sequence of the wild type   |
|                 |             | adenovirus genome, and a gene which encodes an HIV Nef protein (SEQ ID NO: 15)   |
|                 | İ           | incested in F1   |
|                 |             | The claims are directed to a method of making and harvesting of a recombinant  |
| 31              | 76-80       | adenoviral particle that contains a gene encoding an HIV Nef protein.  |
|                 | 2.05        | The claims are directed to a method of generating a cellular mediated immune   |
| 32              | 81, 84, 85  | response to HIV Nef with the recombinant adenoviral particle.  |
|                 |             | The claims are directed to a method of generating a cellular mediated immune   |
| 33              | 82, 83      | response to HIV Nef with the recombinant adenoviral particle in addition to  |
|                 | 1           | response to HIV Net with the recombinant authovital paragraph  |
|                 | ·           | administering a DNA plasmid vaccine.   |
| 34              | 86a         | The claim is drawn to a multivalent vaccine wherein gag, pol and nef are expressed   |
| J <del>-4</del> |             | from three individual vectors.   |
| 26              | 86b, 88, 89 | The claims are drawn to a multivalent vaccine wherein gag, pol and nef are expressed   |
| 35              | 000,        | from one individual vectors.   |
|                 | 86c, 88     | The claims are drawn to a multivalent vaccine wherein gag, pol and nef are expressed   |
| 36              | 300, 00     | from two individual vectors, one expressing nef-pol fusion and one expressing gag.   |
|                 | 86d, 87, 88 | The claims are drawn to a multivalent vaccine wherein gag, pol and nej are expressed   |
| 37              | 800, 87, 80 | from two individual vectors, one expressing gag-pol fusion and one expressing nej.   |
|                 | 1           | The claims are drawn to a multivalent vaccine wherein gag, pol and nej are expressed   |
| 38              | 86e, 88     | from two individual vectors, one expressing nef-gag fusion and one expressing pol.   |
|                 |             | The claims are drawn to a multivalent vaccine wherein gag, pol and nef are expressed   |
| 39              | 86f, 88     | from a single vectors as a fusion protein.   |
|                 |             | The claims are drawn to a multivalent vaccine wherein gag and pol are expressed  |
| 40              | 86g, 88     | from two individual vectors.   |
|                 |             | The claims are drawn to a multivalent vaccine wherein gag and pol are expressed  |
| 41              | 86h, 88, 89 | The claims are trawn to a maintvatch vaccine wherein 608 and per and the first the fir |
|                 |             | individually from one vector.  |
| 42              | 86i, 88     | The claims are drawn to a multivalent vaccine wherein pol and nef are expressed  |
|                 |             | from two individual vectors.   |
| 43              | 86j, 88, 89 | The claims are drawn to a multivalent vaccine wherein pol and nef are expressed  |
| 73              | 1 - 5''     | from individually from one vector.   |
| 44              | 86k, 88     | The claims are drawn to a multivalent vaccine wherein nef and gag are expressed  |
| 44              | 00,         | individually from one vector.  |
| 45              | 861, 88, 89 | The claims are drawn to a multivalent vaccine wherein nef and gag are expressed  |
| 45              | 801, 00, 25 | individually from one vector.  |
| <del></del>     | 86m, 88     | The claims are drawn to a multivalent vaccine wherein gag and pol are expressed as   |
| 46              | 80III, 30   | fusion protein from one vector.  |
|                 | 06- 00      | The claims are drawn to a multivalent vaccine wherein pol and nef are expressed as   |
| 47              | 86n, 88     | fusion protein from one vector.  |
|                 |             | The claims are drawn to a multivalent vaccine wherein nef and gag are expressed as   |
| 48              | 860, 88     | THE CIAMIS ALE WAYER WATER WATER   |
| 1               | i           | fusion protein from one vector.  |

The inventions listed as Groups 1-48 do not relate to a single general inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons:

The technical feature linking groups 1-33 appears to be a recombinant adenoviral vector wherein the adenoviral vector is at least partially deleted in E1 but the vector may contain more deletions, the vector contains wild type sequences including packaging signals partially deleted in E1 but the vector may contain more deletions, the vector contains wild type sequences including packaging signals and a gene encoding a heterologous HIV protein or fragments thereof. Ertl et al. (WO 96/39178) disclose a recombinant adenoviral vector that is deleted in E1 and partially deleted in E3, the remainder of the adenoviral vector contains wild type sequences. The vector additionally contains an insertion of a heterologous protein which includes HIV proteins (see abstract and claims 1 and 5). Therefore, the technical feature linking the inventions of groups 1-45 does not constitute a special technical feature as defined by PCT Rule 13.2, as it does not define a contribution over the prior art.

The special technical feature of the following groups 1-3, 7-15, 19-30 and 34-48 is considered to be the combination of sequences that is disclosed in each group, see individual claim groupings above for the different sequences. The DNA disclosed in each group is trade up of a different sequence having a different structure and different function.

# **REVISED VERSION**

# (19) World Intellectual Property Organization International Bureau





# (43) International Publication Date 21 March 2002 (21.03.2002)

#### PCT

# (10) International Publication Number WO 02/022080 A3

(51) International Patent Classification7:

\_\_\_\_\_

C12N 15/86

- (21) International Application Number: PCT/US01/28861
- (22) International Filing Date:

14 September 2001 (14.09.2001)

(25) Filing Language:

English

- (26) Publication Language:
- English

(30) Priority Data:

60/233,180

15 September 2000 (15.09.2000) U

- (71) Applicant (for all designated States except US): MERCK & CO., INC. [US/US]; 126 East Lincoln Avenue, Rahway, NJ 07065-0907 (US).
- (72) Inventors; and
- (75) Inventors/Applicants (för US only): EMINI, Emilio, A. [US/US]; 126 East Lincoln Avenue, Rahway, NJ 07065-0907 (US). YOUIL, Rima [AU/US]; 126 East Lincoln Avenue, Rahway, NJ 07065-0907 (US). BETT, Andrew, J. [CA/US]; 126 East Lincoln Avenue, Rahway, NJ 07065-0907 (US). CHEN, Ling [US/US]; 126 East Lincoln Avenue, Rahway, NJ 07065-0907 (US). KASLOW, David, C. [US/US]; 126 East Lincoln Avenue, Rahway, NJ 07065-0907 (US). SHIVER, John, W. [US/US]; 126 East Lincoln Avenue, Rahway, NJ 07065-0907 (US). TONER, Timothy, J. [US/US]; 126 East Lincoln Avenue, Rahway, NJ 07065-0907 (US). CASIMIRO, Daniel, R. [PH/US]; 126 East Lincoln Avenue, Rahway, NJ 07065-0907 (US).
- (74) Common Representative: MERCK & CO., INC.; 126 East Lincoln Avenue, Rahway, NJ 07065-0907 (US).

- (81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.
- (84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

#### Published:

- with international search report
- (88) Date of publication of the international search report:

  2 May 2002

  Date of publication of the revised international search report:

  16 January 2003
- (15) Information about Corrections:

see PCT Gazette No. 03/2003 of 16 January 2003, Section II

**Previous Correction:** 

see PCT Gazette No. 30/2002 of 25 July 2002, Section II

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: ENHANCED FIRST GENERATION ADENOVIRUS VACCINES EXPRESSING CODON OPTIMIZED HIV1-GAG, POL, NEF AND MODIFICATIONS

(57) Abstract: First generation adenoviral vectors and associated recombinant adenovirus-based HIV vaccines which show enhanced stability and growth properties and greater cellular-mediated immunity are described within this specification. These adenoviral vectors are utilized to generate and produce through cell culture various adenoviral-based HIV-1 vaccines which contain HIV-1 gag, HIV-1 pol and/or HIV-1 nef polynucleotide pharmaceutical products, and biologically relevant modifications thereof. These adenovirus vaccines, when directly introduced into living vertebrate tissue, preferably a mammalian host such as a human or a non-human mammal of commercial or domestic veterinary importance, express the HIV1- Gag, Pol and/or Nef protein or biologically modification thereof, inducing a cellular immune response which specifically recognizes HIV-1. The exemplified polynucleotides of the present invention are synthetic DNA molecules encoding HIV-1 Gag, encoding codon optimized HIV-1 Pol, derivatives of optimized HIV-1 Pol (including constructs wherein protease, reverse transcriptase, RNAse H and integrase activity of HIV-1 Pol is inactivated), HIV-1 Nef and derivatives of optimized HIV-1 Nef, including nef mutants which effect wild type characteristics of Nef, such as myristylation and down regulation of host CD4. The adenoviral vaccines of the present invention, when administered alone or in a combined modality regime, will offer a prophylactic advantage to previously uninfected individuals and/or provide a therapeutic effect by reducing viral load levels within an infected individual, thus prolonging the asymptomatic phase of HIV-1 infection.

International application No.

PCT/US01/28861

|  | PC1/0301/28801  |  |  |  |  |
|--|---|--|--|--|--|
| A. CLASSIFICATION OF SUBJECT MATTER  |   |  |  |  |  |
| IPC(7) : C12N 15/86  |   |  |  |  |  |
| US CL : 435/456  |   |  |  |  |  |
| According to International Patent Classification (IPC) or to both a  B. FIELDS SEARCHED  | national classification and IPC   |  |  |  |  |
|  |   |  |  |  |  |
| Minimum documentation searched (classification system followed by classification symbols) U.S.: 424/205.1, 207.1, 227.1, 233.1; 435/69.1, 69.3, 173.3, 235.1, 320.1, 456; 530/23.72; |   |  |  |  |  |
| Demonstration and the Australia Australia  |   |  |  |  |  |
| Documentation searched other than minimum documentation to the   | e extent that such documents are included in the helds searched   |  |  |  |  |
|  |   |  |  |  |  |
|  |   |  |  |  |  |
| Electronic data base consulted during the international search (nat Please See Continuation Sheet  | me of data base and, where practicable, search terms used)  |  |  |  |  |
|  |   |  |  |  |  |
| C. DOCUMENTS CONSIDERED TO BE RELEVANT   |   |  |  |  |  |
| Category * Citation of document, with indication, where a  | ppropriate, of the relevant passages Relevant to claim No.  |  |  |  |  |
| X WO 96/39178 (ERTL et al.) 12 December 1996 (13   | 2.12.1996), see page 5, 6,10, 12, 13 1-3, 8-11, 18  |  |  |  |  |
| and claims 1 and 5.  |   |  |  |  |  |
| Y  | 4, 5, 13-17, 29-32, 34, 35, 37  |  |  |  |  |
| X US 6,019,978 A (ERTL et al.) 1 February 2000,(0  | 1/02/2000), see columns 2, 7 and 8.   |  |  |  |  |
| Y  | 4, 5, 13-17, 29-32, 34, 35, 37  |  |  |  |  |
| X,P US 6,287,571 B/ (ERTL et al.) 11 September 200 and claim 1.  | 01 (11/09/2001), see columns 2, 7, 8 1, 9, 18   |  |  |  |  |
| X US 5,643,579A (HUNG et al.) 1 July 1997 (01/07/  | /1997), see examples 1, 2, 25 and 26.   |  |  |  |  |
| Y  | 4,5,13-17, 29-32, 34,<br>35, 37   |  |  |  |  |
| Y WANG et al. The use of an E1-deleted, replication expressing the rabies virus glycoprotein for early v   | raccination of mice against rabies virus.   |  |  |  |  |
| Journal of Virology (March 1997) Vol. 71, No. 5,   | pp 3677-3683.   |  |  |  |  |
|  | <u> </u>  |  |  |  |  |
|  |   |  |  |  |  |
| Further documents are listed in the continuation of Box C.   | See patent family annex.  |  |  |  |  |
| Special categories of cited documents:   | T later document published after the international filing date or   |  |  |  |  |
| "A" document defining the general state of the art which is not considered to be of particular relevance   | priority date and not in conflict with the application but cited to<br>understand the principle or theory underlying the invention  |  |  |  |  |
| "E" earlier application or patent published on or after the international filing date  | "X" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is taken alone  |  |  |  |  |
| "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)              | "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such |  |  |  |  |
| combination being obvious to a person skilled in the art  O' document referring to an oral disclosure, use, exhibition or other means  |   |  |  |  |  |
| *A" document member of the same patent family  "P" document published prior to the international filing date but later than the  |   |  |  |  |  |
| Date of the actual completion of the international search  | Date of mailing of the international search report  |  |  |  |  |
| 06 February 2002 (06.02.2002) 19 AUG 2002  |   |  |  |  |  |
| Name and mailing address of the ISA/US   | Authorized officer To 10 x blling La  |  |  |  |  |
| Commissioner of Patents and Trademarks Box PCT   | Ulrike Winkler, Ph.D. Nell State of the   |  |  |  |  |
| Washington, D.C. 20231   | 1)  |  |  |  |  |
| Facsimile No. (703)305-3230 Telephone No. 703-308-0196   |   |  |  |  |  |

Form PCT/ISA/210 (second sheet) (July 1998)

International application No.

PCT/US01/28861

|          | Citation of document, with indication, where appropriate, of the relevant passages  | Relevant to claim No. |
|----------|---|-----------------------|
| <b>Y</b> | NATUK et al. Immunogenicity of recombinant human adenovirus -human immunodeficiency virus vaccines in chimpanzees. Aids Research and Human Retroviruses (1993) Vol. 9, No. 5, pp395-404, see material and methods.  | 1, 9, 29-32           |
| Y        | PREVEC et al. Immune response to HIV-1 gag antigens induced by recombinant adenovirus vectors in mice and rhesus macaque monkeys. Journal of Acquired Immune Deficincy Syndrome. (1991) Vol. 4, No. 6 pp. 568-76, see abstract.                           | 1, 9, 29-32           |
| Y        | LORI et al. Rapid protection against human immunodeficiency virus type 1 (HIV-1) replication mediated by high efficiency non-retroviral delivery of genes interfering with HIV-1 tat and gag. Gene Therapy (1994) Vol. 1, No. 1, pp. 27-31, see abstract. | 1, 9                  |
| Y        | PFARR et al. Differential effects of polyadenylation regions on gene expression in mammalian cells. DNA (1986) Vol. 5, No. 2, pp.115-22, see abstract.  | 16                    |
| Y        | NATUK et al. Adenovirus vectored vaccine. Developmental Biological Standards (1994)<br>Vol. 82, pp. 71-77, see abstract.  | 1, 9                  |
|          |   | ٠.                    |
|          |   |                       |
|          |   |                       |
|          |   | •                     |
|          |   |                       |
|          |   |                       |
|          |   |                       |
|          |   | ·                     |
| •        |   | -                     |
|          | ·   |                       |
|          |   |                       |
|          |   |                       |
|          | ·   |                       |

International application No.

PCT/US01/28861

| Box I Observations where certain claims were found unsearchable (Continuation of Item 1 of first sheet)   |  |  |  |  |
|---|--|--|--|--|
| This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:   |  |  |  |  |
| Claim Nos.:     because they relate to subject matter not required to be searched by this Authority, namely:  |  |  |  |  |
| Claim Nos.:  because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:                              |  |  |  |  |
| 3. Claim Nos.:  because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).   |  |  |  |  |
| Box II Observations where unity of invention is lacking (Continuation of Item 2 of first sheet)   |  |  |  |  |
| This International Searching Authority found multiple inventions in this international application, as follows:  Please See Continuation Sheet  |  |  |  |  |
| As all required additional search fees were timely paid by the applicant, this international search report covers all   |  |  |  |  |
| searchable claims.  2. As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite  |  |  |  |  |
| payment of any additional fee.  3. As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:                 |  |  |  |  |
|   |  |  |  |  |
| 4. No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.: 1-5, 8-11, 13-18, 29-32, 34, 35, 37 |  |  |  |  |
| Remark on Protest  The additional search fees were accompanied by the applicant's protest.  No protest accompanied the payment of additional search fees.   |  |  |  |  |

Form PCT/ISA/210 (continuation of first sheet(1)) (July 1998)

International application No.

PCT/US01/28861

BOX II. OBSERVATIONS WHERE UNITY OF INVENTION IS LACKING

This application contains the following inventions or groups of inventions which are not so linked as to form a single general inventive concept under PCT Rule 13.1. In order for all inventions to be examined, the appropriate additional examination fees must be paid.

| Group | Claims  |  |
|-------|---|--|
| 1     | 1-5, 8-11,<br>13-18, 29,<br>30, 31, 32,<br>34, 35, 37 | The claims are directed to an adenoviral vector that is at least partially deleted of<br><u>AEI</u> , the vector contains the cis-acting packaging sequence of the wild type<br>adenovirus genome, and a gene which encodes an <u>HIV Gag protein (SEQ ID NO: 29)</u><br>inserted in the <u>parallel orientation of E1</u> . In addition the vector contains a promoter<br>and a polyadenylation signal. |
| 2     | 6, 7, 36  | The claims are directed to an adenoviral vector that is at least partially deleted of $\Delta E1$ and $\Delta E3$ , the vector contains the cis-acting packaging sequence of the wild type adenovirus genome, and a gene which encodes an HIV Gag protein (SEQ ID NO: 29).   |
| 3     | 12, 33  | The claims are directed to an adenoviral vector that is at least partially deleted of $\Delta E1$ , the vector contains the cis-acting packaging sequence of the wild type adenovirus genome, and a gene which encodes an HIV protein inserted in the antiparallel orientation of E1.  |
| 4     | 19-23, 38-42  | The claims are directed to a method of making and harvesting of a recombinant adenoviral particle that contains a gene encoding an HIV Gag protein.  |
| 5     | 24, 27, 28,<br>43, 46, 47                             | The claim is directed to a method of generating a cellular mediated immune response to HIV Gag protein with the recombinant adenoviral particle.   |
| 6     | 25, 26, 44,<br>45                                     | The claim is directed to a method of generating a cellular mediated immune response to HIV Gag protein with the recombinant adenoviral particle in addition to administering a DNA plasmid vaccine.  |
| 7     | 48-51, 53,<br>54, 56                                  | The claims are directed to an adenoviral vector that is at least partially deleted of $\Delta E1$ , the vector contains the cis-acting packaging sequence of the wild type adenovirus genome, and a gene which encodes an HIV Pol protein (SEQ ID NO: 1) inserted in the parallel orientation of E1.   |
| 8     | 48-51, 53,<br>54, 56                                  | The claims are directed to an adenoviral vector that is at least partially deleted of <a href="MEL"><u>AEI</u></a> , the vector contains the cis-acting packaging sequence of the wild type adenovirus genome, and a gene which encodes an <a href="HIV Pol protein (SEQ ID NO: 5">HIV Pol protein (SEQ ID NO: 5)</a> inserted in the parallel orientation of E1.  |
| 9     | 48-51, 53,<br>54, 56                                  | The claims are directed to an adenoviral vector that is at least partially deleted of $\Delta E1$ , the vector contains the cis-acting packaging sequence of the wild type adenovirus genome, and a gene which encodes an HIV Pol protein (SEQ ID NO: 7) inserted in the parallel orientation of E1.   |
| 10    | 52  | The claim is directed to an adenoviral vector that is at least partially deleted of $\Delta E1$ , the vector contains the cis-acting packaging sequence of the wild type adenovirus genome, and a gene which encodes an HIV Pol protein (SEQ ID NO: 1) inserted in the antiparallel orientation of E1.   |
| 11    | 52  | The claim is directed to an adenoviral vector that is at least partially deleted of $\Delta E1$ , the vector contains the cis-acting packaging sequence of the wild type adenovirus genome, and a gene which encodes an HIV Pol protein (SEQ ID NO: 5) inserted in the antiparallel orientation of E1.   |
| 12    | 52  | The claim is directed to an adenoviral vector that is at least partially deleted of $\Delta E1$ , the vector contains the cis-acting packaging sequence of the wild type adenovirus genome, and a gene which encodes an HIV Pol protein (SEQ ID NO: 7) inserted in the antiparallel orientation of E1.   |
| 13    | 55  | The claim is directed to an adenoviral vector that is at least partially deleted of $\Delta E_1$   |

Form PCT/ISA/210 (second sheet) (July 1998)

International application No.

PCT/US01/28861

|     | _ •           |  |
|-----|---------------|--|
|     |               |  |
|     | ł             | and ΔE3, the vector contains the cis-acting packaging sequence of the wild type  |
|     | -             | adenovirus genome, and a gene which encodes an HIV Pol protein (SEQ ID NO: 1) inserted in E1.  |
| 14  | 55            | The claim is directed to an adenoviral vector that is at least partially deleted of $\Delta E1$  |
|     |               | and $\Delta E3$ , the vector contains the cis-acting packaging sequence of the wild type   |
|     | 1             | adenovirus genome, and a gene which encodes an HIV Pol protein (SEQ ID NO: 5) inserted in E1.  |
| 15  | 55            | The claim is directed to an adenoviral vector that is at least partially deleted of $\Delta E1$  |
| 15  | 33            | and AE3, the vector contains the cis-acting packaging sequence of the wild type  |
|     |               | adenovirus genome, and a gene which encodes an HIV Pol protein (SEQ ID NO: 7)  |
| ·   | <u> </u>      | inserted in E1.  |
| 16  | 57-61         | The claims are directed to a method of making and harvesting of a recombinant  |
|     |               | adenoviral particle that contains a gene encoding an HIV Pol protein.  |
| 17  | 62, 65, 66    | The claim is directed to a method of generating a cellular mediated immune response  |
| 18  | 63, 64        | to HIV Pol protein with the recombinant adenoviral particle.  The claim is directed to a method of generating a cellular mediated immune response  |
| 10  | 05,04         | to HIV Pol protein with the recombinant adenoviral particle in addition to   |
|     |               | administering a DNA plasmid vaccine.   |
| 19  | 67-70, 72,    | The claims are directed to an adenoviral vector that is at least partially deleted of  |
| .,  | 73, 75        | AE1, the vector contains the cis-acting packaging sequence of the wild type  |
|     | 1.5,.5        | adenovirus genome, and a gene which encodes an HIV Nef protein (SEQ ID NO: 9)  |
|     | 1             | inserted in the parallel orientation of E1.  |
| 20  | 67-70, 72,    | The claims are directed to an adenoviral vector that is at least partially deleted of  |
|     | 73, 75        | $\Delta E1$ , the vector contains the cis-acting packaging sequence of the wild type   |
|     |               | adenovirus genome, and a gene which encodes an HIV Nef protein (SEQ ID NO: 11)   |
|     | Î             | inserted in the parallel orientation of E1.  |
| 21  | 67-70, 72,    | The claims are directed to an adenoviral vector that is at least partially deleted of  |
|     | 73, 75        | $\Delta E1$ , the vector contains the cis-acting packaging sequence of the wild type   |
|     | ı             | adenovirus genome, and a gene which encodes an HIV Nef protein (SEQ ID NO: 13)   |
|     |               | inserted in the parallel orientation of E1.  |
| 22  | 67-70, 72,    | The claims are directed to an adenoviral vector that is at least partially deleted of  |
|     | 73, 75        | $\Delta E1$ , the vector contains the cis-acting packaging sequence of the wild type   |
|     | İ             | adenovirus genome, and a gene which encodes an HIV Nef protein (SEQ ID NO: 15)   |
|     |               | inserted in the parallel orientation of E1.  |
| 23  | 71            | The claim is directed to an adenoviral vector that is at least partially deleted of $\Delta E1$ ,  |
|     | 1             | the vector contains the cis-acting packaging sequence of the wild type adenovirus  |
|     | j             | genome, and a gene which encodes an HIV Nef protein (SEQ ID NO: 9) inserted in the antiparallel orientation of E1.   |
| 24  | 71            | The claim is directed to an adenoviral vector that is at least partially deleted of $\Delta E1$ ,  |
| 2.4 |               | the vector contains the cis-acting packaging sequence of the wild type adenovirus  |
|     | 1             | genome, and a gene which encodes an HIV Nef protein (SEQ ID NO: 11) inserted in  |
|     |               | the antiparallel orientation of E1.  |
| 25  | 71            | The claim is directed to an adenoviral vector that is at least partially deleted of $\Delta E1$ ,  |
|     |               | the vector contains the cis-acting packaging sequence of the wild type adenovirus  |
|     | ]             | genome, and a gene which encodes an HIV Nef protein (SEQ ID NO: 13) inserted in  |
|     |               | the antiparallel orientation of E1.  |
| 26  | 71            | The claim is directed to an adenoviral vector that is at least partially deleted of $\Delta E1$ .  |
|     |               | the vector contains the cis-acting packaging sequence of the wild type adenovirus  |
|     |               | genome, and a gene which encodes an HIV Nef protein (SEQ ID NO: 15) inserted in  |
|     |               | the antiparallel orientation of E1.  |
| 27  | 74            | The claim is directed to an adenoviral vector that is at least partially deleted of $\Delta E_1$   |
|     | ]             | and $\Delta E3$ , the vector contains the cis-acting packaging sequence of the wild type   |
|     | 1             | adenovirus genome, and a gene which encodes an HIV Nef protein (SEQ ID NO: 9)  |
| -   | <del>- </del> | inserted in E1.  |
| 28  | 74            | The claim is directed to an adenoviral vector that is at least partially deleted of $\Delta E_1$   |
|     |               | and $\Delta E3$ , the vector contains the cis-acting packaging sequence of the wild type   |
|     | 1             | I administrate management and a constability of the control of the |
|     |               | adenovirus genome, and a gene which encodes an HIV Nef protein (SEQ ID NO: 11)   |
| 20  | 74            | inserted in E1.  |
| 29  | 74            |  |

International application No.

PCT/US01/28861

|    |             | adenovirus genome, and a gene which encodes an HIV Nef protein (SEQ ID NO: 13) inserted in E1.   |  |
|----|-------------|--|--|
| 30 | 74          | The claim is directed to an adenoviral vector that is at least partially deleted of <u>AE1</u> and <u>AE3</u> , the vector contains the cis-acting packaging sequence of the wild type adenovirus genome, and a gene which encodes an <u>HIV Nef protein (SEQ ID NO: 15)</u> inserted in E1. |  |
| 31 | 76-80       | The claims are directed to a method of making and harvesting of a recombinant adenoviral particle that contains a gene encoding an HIV Net protein.  |  |
| 32 | 81, 84, 85  | The claims are directed to a method of generating a cellular mediated immune response to HIV Nef with the recombinant adenoviral particle.   |  |
| 33 | 82, 83      | The claims are directed to a method of generating a cellular mediated immune response to HIV Nef with the recombinant adenoviral particle in addition to administering a DNA plasmid vaccine.  |  |
| 34 | 86a         | The claim is drawn to a multivalent vaccine wherein gag, pol and nef are expressed from three individual vectors.  |  |
| 35 | 86b, 88, 89 | The claims are drawn to a multivalent vaccine wherein gag, pol and nef are expressed from one individual vectors.  |  |
| 36 | 86c, 88     | The claims are drawn to a multivalent vaccine wherein gag, pol and nef are expressed from two individual vectors, one expressing nef-pol fusion and one expressing gag.  |  |
| 37 | 86d, 87, 88 | The claims are drawn to a multivalent vaccine wherein gag, pol and nef are expressed from two individual vectors, one expressing gag-pol fusion and one expressing nef.  |  |
| 38 | 86e, 88     | The claims are drawn to a multivalent vaccine wherein gag, pol and nef are expressed from two individual vectors, one expressing nef-gag fusion and one expressing pol.  |  |
| 39 | 86f, 88     | The claims are drawn to a multivalent vaccine wherein gag, pol and nef are expressed from a single vectors as a fusion protein.  |  |
| 40 | 86g, 88     | The claims are drawn to a multivalent vaccine wherein gag and pol are expressed from two individual vectors.   |  |
| 41 | 861, 88, 89 | The claims are drawn to a multivalent vaccine wherein gag and pol are expressed individually from one vector.  |  |
| 42 | 86i, 88     | The claims are drawn to a multivalent vaccine wherein pol and nef are expressed from two individual vectors.   |  |
| 43 | 86j, 88, 89 | The claims are drawn to a multivalent vaccine wherein pol and nef are expressed from individually from one vector.   |  |
| 44 | 86k, 88     | The claims are drawn to a multivalent vaccine wherein nef and gag are expressed individually from one vector.  |  |
| 45 | 861, 88, 89 | The claims are drawn to a multivalent vaccine wherein nef and gag are expressed individually from one vector.  |  |
| 46 | 86m, 88     | The claims are drawn to a multivalent vaccine wherein gag and pol are expressed as a fusion protein from one vector.   |  |
| 47 | 86n, 88     | The claims are drawn to a multivalent vaccine wherein pol and nef are expressed as a fusion protein from one vector.   |  |
|    |             |  |  |

The inventions listed as Groups 1-48 do not relate to a single general inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons:

The technical feature linking groups 1-33 appears to be a recombinant adenoviral vector wherein the adenoviral vector is at least partially deleted in E1 but the vector may contain more deletions, the vector contains wild type sequences including packaging signals and a gene encoding a heterologous HIV protein or fragments thereof. Ertl et al. (WO 96/39178) disclose a recombinant adenoviral vector that is deleted in E1 and partially deleted in E3, the remainder of the adenoviral vector contains wild type sequences. The vector additionally contains an insertion of a heterologous protein which includes HIV proteins (see abstract and claims 1 and 5). Therefore, the technical feature linking the inventions of groups 1-45 does not constitute a special technical feature as defined by PCT Rule 13.2, as it does not define a contribution over the prior art.

The special technical feature of the following groups 1-3, 7-15, 19-30 and 34-48 is considered to be the combination of sequences that is disclosed in each group, see individual claim groupings above for the different sequences. The DNA disclosed in each group is made up of a different sequence having a different structure and different function.

International application No.

PCT/US01/28861

The special technical feature of group 4, 16 and 31 is considered to be a method of producing recombinant adenoviral particles. Each group contains different sequences hence the resulting particles would have different structures and functions associated with the

The special technical feature of group 5, 17 and 32 is considered to be a method of producing a cellular mediated immune response to the heterologous protein encoded by the different adenoviral vectors. Each group contains different sequences a encoding different protein, therefore the resulting immune response will also be different.

The special technical feature of group 6, 18 and 33 is considered to be a method of producing a cellular mediated immune response to the heterologous protein encoded by the different adenoviral vectors in conjunction with immunizing the individual a DNA plasmid vaccine. Each method contains different sequences encoding a different protein, therefore the resulting immune response will also be different.

Accordingly, groups 1-48 are not so linked by the same or corresponding technical feature as to form a single general inventive concept.

Continuation of B. FIELDS SEARCHED Item 3:

WEST 2.0, STN-BIOSIS, MEDLINE

adenoviral vector, deletion, HIV, Gag, polyadenylation signal, CMV promoter

Form PCT/ISA/210 (second sheet) (July 1998)

# CORRECTED VERSION

# (19) World Intellectual Property Organization International Bureau



# 

# (43) International Publication Date 21 March 2002 (21.03.2002)

### PCT

# (10) International Publication Number WO 02/022080 A3

- (51) International Patent Classification7: C12N 15/86
- (21) International Application Number: PCT/US01/28861
- (22) International Filing Date:

14 September 2001 (14.09.2001)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data: 60/233,180 15 September 2

15 September 2000 (15.09.2000) US

- (71) Applicant (for all designated States except US): MERCK & CO., INC. [US/US]; 126 East Lincoln Avenue, Rahway, NJ 07065-0907 (US).
- (72) Inventors; and
- (75) Inventors/Applicants (for US only): EMINI, Emilio, A. [US/US]; 126 East Lincoln Avenue, Rahway, NJ 07065-0907 (US). YOUIL, Rima [AU/US]; 126 East Lincoln Avenue, Rahway, NJ 07065-0907 (US). BETT, Andrew, J. [CA/US]; 126 East Lincoln Avenue, Rahway, NJ 07065-0907 (US). CHEN, Ling [US/US]; 126 East Lincoln Avenue, Rahway, NJ 07065-0907 (US). KASLOW, David, C. [US/US]; 126 East Lincoln Avenue, Rahway, NJ 07065-0907 (US). SHIVER, John, W. [US/US]; 126 East Lincoln Avenue, Rahway, NJ 07065-0907 (US). TONER, Timothy, J. [US/US]; 126 East Lincoln Avenue, Rahway, NJ 07065-0907 (US). CASIMIRO, Daniel, R. [PH/US]; 126 East Lincoln Avenue, Rahway, NJ 07065-0907 (US).
- (74) Common Representative: MERCK & CO., INC.; 126 East Lincoln Avenue, Rahway, NJ 07065-0907 (US).

- (81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.
- (84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

#### Published:

- with international search report
- (88) Date of publication of the international search report:

  2 May 2002

  Date of publication of the revised international search report:

  16 January 2003
- (48) Date of publication of this corrected version:

6 March 2003

(15) Information about Corrections:

see PCT Gazette No. 10/2003 of 6 March 2003, Section II Previous Corrections:

see PCT Gazette No. 03/2003 of 16 January 2003, Section II

see PCT Gazette No. 30/2002 of 25 July 2002, Section II

[Continued on next page]

(54) Title: ENHANCED FIRST GENERATION ADENOVIRUS VACCINES EXPRESSING CODON OPTIMIZED HIV1-GAG, POL, NEF AND MODIFICATIONS

(57) Abstract: First generation adenoviral vectors and associated recombinant adenovirus-based HIV vaccines which show enhanced stability and growth properties and greater cellular-mediated immunity are described within this specification. These adenoviral vectors are utilized to generate and produce through cell culture various adenoviral-based HIV-1 vaccines which contain HIV-1 gag, HIV-1 pol and/or HIV-1 nef polynucleotide pharmaceutical products, and biologically relevant modifications thereof. These adenovirus vaccines, when directly introduced into living vertebrate tissue, preferably a mammalian host such as a human or a non-human mammal of commercial or domestic veterinary importance, express the HIV1-Gag, Pol and/or Nef protein or biologically modification thereof, inducing a cellular immune response which specifically recognizes HIV-1. The exemplified polynucleotides of the present invention are synthetic DNA molecules encoding HIV-1 Gag, encoding codon optimized HIV-1 Pol, derivatives of optimized HIV-1 Pol (including constructs wherein protease, reverse transcriptase, RNAse H and integrase activity of HIV-1 Pol is inactivated), HIV-1 Nef and derivatives of optimized HIV-1 Nef, including nef mutants which effect wild type characteristics of Nef, such as myristylation and down regulation of host CD4. The adenoviral vaccines of the present invention, when administered alone or in a combined modality regime, will offer a prophylactic advantage to previously uninfected individuals and/or provide a therapeutic effect by reducing viral load levels within an infected individual, thus prolonging the asymptomatic phase of HIV-1 infection.





For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

#### TITLE OF THE INVENTION

ENHANCED FIRST GENERATION ADENOVIRUS VACCINES EXPRESSING CODON OPTIMIZED HIV1-GAG, POL, NEF AND MODIFICATIONS

### 5 CROSS-REFERENCE TO RELATED APPLICATIONS

This application claims the benefit, under 35 U.S.C. §119(e), of U.S. provisional applications 60/233,180, 60/279,056, and Attorney Docket 20867PV2 (serial number unassigned), filed September 15, 2000, March 27, 2001, and September 7, 2001, respectively.

10

15

20

25

30

35

STATEMENT REGARDING FEDERALLY-SPONSORED R&D Not Applicable

# REFERENCE TO MICROFICHE APPENDIX

Not Applicable

#### FIELD OF THE INVENTION

The present invention relates to recombinant, replication-deficient first generation adenovirus vaccines found to exhibit enhanced growth properties and greater cellular-mediated immunity as compared to other replication-deficient vectors. The invention also relates to the associated first generation adenoviral vectors described herein, which, through the incorporation of additional 5' adenovirus sequence, enhance large scale production efficiency of the recombinant, replicationdefective adenovirus described herein. Another aspect of the instant invention is the surprising discovery that the intron A portion of the human cytomegalovirus (hCMV) promoter constitutes a region of instability in adenoviral vector constructs. Removal of this region from adenoviral expression constructs results in greatly improved vector stability. Therefore, improved vectors expressing a transgene under the control of an intron A-deleted CMV promoter constitute a further aspect of this invention. These adenoviral vectors are useful for generating recombinant adenovirus vaccines against human immunodeficiency virus (HIV). In particular, the first generation adenovirus vectors disclosed herein are utilized to construct and generate adenovirus-based HIV-1 vaccines which contain HTV-1 Gag, HTV-1 Pol and/or HTV-1 Nef polynucleotide pharmaceutical products, and biologically active modifications thereof. Host administration of the recombinant, replication-deficient adenovirus vaccines described herein results in expression of HIV-1 Gag, HIV-1- Pol and/or Nef protein or

immunologically relevant modifications thereof, inducing a cellular immune response which specifically recognizes HIV-1. The exemplified polynucleotides of the present invention are synthetic DNA molecules encoding codon optimized HIV-1 Gag, HIV-1 Pol, derivatives of optimized HIV-1 Pol (including constructs wherein protease, reverse transcriptase, RNAse H and integrase activity of HIV-1 Pol is inactivated), HIV-1 Nef, and derivatives of optimized HIV-1 Nef, including nef mutants which effect wild type characteristics of Nef, such as myristylation and down regulation of host CD4. The HIV adenovirus vaccines of the present invention, when administered alone or in a combined modality and/or prime/boost regimen, will offer a prophylactic advantage to previously uninfected individuals and/or provide a therapeutic effect by reducing viral load levels within an infected individual, thus prolonging the asymptomatic phase of HIV-1 infection.

#### BACKGROUND OF THE INVENTION

10

15

20

25

30

Human Immunodeficiency Virus-1 (HIV-1) is the etiological agent of acquired human immune deficiency syndrome (AIDS) and related disorders. HIV-1 is an RNA virus of the Retroviridae family and exhibits the 5'LTR-gag-pol-env-LTR 3' organization of all retroviruses. The integrated form of HIV-1, known as the provirus, is approximately 9.8 Kb in length. Each end of the viral genome contains flanking sequences known as long terminal repeats (LTRs). The HIV genes encode at least nine proteins and are divided into three classes; the major structural proteins (Gag, Pol, and Env), the regulatory proteins (Tat and Rev); and the accessory proteins (Vpu, Vpr, Vif and Nef).

The gag gene encodes a 55-kilodalton (kDa) precursor protein (p55) which is expressed from the unspliced viral mRNA and is proteolytically processed by the HIV protease, a product of the pol gene. The mature p55 protein products are p17 (matrix), p24 (capsid), p9 (nucleocapsid) and p6.

The pol gene encodes proteins necessary for virus replication; a reverse transcriptase, a protease, integrase and RNAse H. These viral proteins are expressed as a Gag-Pol fusion protein, a 160 kDa precursor protein which is generated via a ribosomal frame shifting. The viral encoded protease proteolytically cleaves the Pol polypeptide away from the Gag-Pol fusion and further cleaves the Pol polypeptide to the mature proteins which provide protease (Pro, P10), reverse transcriptase (RT, P50), integrase (IN, p31) and RNAse H (RNAse, p15) activities.

The *nef* gene encodes an early accessory HIV protein (Nef) which has been shown to possess several activities such as down regulating CD4 expression, disturbing T-cell activation and stimulating HIV infectivity.

The *env* gene encodes the viral envelope glycoprotein that is translated as a 160-kilodalton (kDa) precursor (gp160) and then cleaved by a cellular protease to yield the external 120-kDa envelope glycoprotein (gp120) and the transmembrane 41-kDa envelope glycoprotein (gp41). Gp120 and gp41 remain associated and are displayed on the viral particles and the surface of HIV-infected cells.

5

10

15

20

25

30

35

The *tat* gene encodes a long form and a short form of the Tat protein, a RNA binding protein which is a transcriptional transactivator essential for HIV-1 replication.

The *rev* gene encodes the 13 kDa Rev protein, a RNA binding protein. The Rev protein binds to a region of the viral RNA termed the Rev response element (RRE). The Rev protein promotes transfer of unspliced viral RNA from the nucleus to the cytoplasm. The Rev protein is required for HIV late gene expression and in turn, HIV replication.

- Gp120 binds to the CD4/chemokine receptor present on the surface of helper T-lymphocytes, macrophages and other target cells in addition to other co-receptor molecules. X4 (macrophage tropic) virus show tropism for CD4/CXCR4 complexes while a R5 (T-cell line tropic) virus interacts with a CD4/CCR5 receptor complex. After gp120 binds to CD4, gp41 mediates the fusion event responsible for virus entry. The virus fuses with and enters the target cell, followed by reverse transcription of its single stranded RNA genome into the double-stranded DNA via a RNA dependent DNA polymerase. The viral DNA, known as provirus, enters the cell nucleus, where the viral DNA directs the production of new viral RNA within the nucleus, expression of early and late HIV viral proteins, and subsequently the production and cellular release of new virus particles. Recent advances in the ability to detect viral load within the host shows that the primary infection results in an extremely high generation and tissue distribution of the virus, followed by a steady state level of virus (albeit through a continual viral production and turnover during this phase), leading ultimately to another burst of virus load which leads to the onset of clinical AIDS. Productively infected cells have a half life of several days, whereas chronically or latently infected cells have a 3-week half life, followed by non-productively infected cells which have a long half life (over 100 days) but do not significantly contribute to day to day viral loads seen throughout the course of disease.

Destruction of CD4 helper T lymphocytes, which are critical to immune defense, is a major cause of the progressive immune dysfunction that is the hallmark of HIV infection. The loss of CD4 T-cells seriously impairs the body's ability to fight most invaders, but it has a particularly severe impact on the defenses against viruses, fungi, parasites and certain bacteria, including mycobacteria.

5

10

15

20

25

30

35

Effective treatment regimens for HTV-1 infected individuals have become available recently. However, these drugs will not have a significant impact on the disease in many parts of the world and they will have a minimal impact in halting the spread of infection within the human population. As is true of many other infectious diseases, a significant epidemiologic impact on the spread of HIV-1 infection will only occur subsequent to the development and introduction of an effective vaccine. There are a number of factors that have contributed to the lack of successful vaccine development to date. As noted above, it is now apparent that in a chronically infected person there exists constant virus production in spite of the presence of anti-HIV-1 humoral and cellular immune responses and destruction of virally infected cells. As in the case of other infectious diseases, the outcome of disease is the result of a balance between the kinetics and the magnitude of the immune response and the pathogen replicative rate and accessibility to the immune response. Pre-existing immunity may be more successful with an acute infection than an evolving immune response can be with an established infection. A second factor is the considerable genetic variability of the virus. Although anti-HIV-1 antibodies exist that can neutralize HIV-1 infectivity in cell culture, these antibodies are generally virus isolate-specific in their activity. It has proven impossible to define serological groupings of HIV-1 using traditional methods. Rather, the virus seems to define a serological "continuum" so that individual neutralizing antibody responses, at best, are effective against only a handful of viral variants. Given this latter observation, it would be useful to identify immunogens and related delivery technologies that are likely to elicit anti-HTV-1 cellular immune responses. It is known that in order to generate CTL responses antigen must be synthesized within or introduced into cells, subsequently processed into small peptides by the proteasome complex, and translocated into the endoplasmic reticulum/Golgi complex secretory pathway for eventual association with major histocompatibility complex (MHC) class I proteins. CD8<sup>+</sup> T lymphocytes recognize antigen in association with class I MHC via the T cell receptor (TCR) and the CD8 cell surface protein. Activation of naive CD8<sup>+</sup> T cells into activated effector or memory cells generally requires both TCR engagement of antigen as described above as well as engagement of costimulatory proteins. Optimal

induction of CTL responses usually requires "help" in the form of cytokines from CD4<sup>+</sup> T lymphocytes which recognize antigen associated with MHC class II molecules via TCR and CD4 engagement.

5

10

15

20

25

30

35

European Patent Applications 0 638 316 (Published February 15, 1995) and 0 586 076 (Published March 9, 1994), (both assigned to American Home Products Corporation) describe replicating adenovirus vectors carrying an HIV gene, including env or gag. Various treatment regimens were used with chimpanzees and dogs, some of which included booster adenovirus or protein plus alum treatments.

Replication-defective adenoviral vectors harboring deletions in the E1 region are known, and recent adenoviral vectors have incorporated the known packaging repeats into these vectors; e.g., see EP 0 707 071, disclosing, *inter alia*, an adenoviral vector deleted of E1 sequences from base pairs 459 to 3328; and U.S. Patent No. 6,033,908, disclosing, *inter alia*, an adenoviral vector deleted of base pairs 459-3510. The packaging efficiency of adenovirus has been taught to depend on the number of incorporated individual A (packaging) repeats; *see*, *e.g.*, Gräble and Hearing, 1990 *J. Virol.* 64(5):2047-2056; Gräble and Hearing, 1992 *J. Virol.* 66(2):723-731.

Larder, et al., (1987, Nature 327: 716-717) and Larder, et al., (1989, Proc. Natl. Acad. Sci. 86: 4803-4807) disclose site specific mutagenesis of HIV-1 RT and the effect such changes have on in vitro activity and infectivity related to interaction with known inhibitors of RT.

Davies, et al. (1991, *Science* 252:, 88-95) disclose the crystal structure of the RNase H domain of HIV-1 Pol.

Schatz, et al. (1989, FEBS Lett. 257: 311-314) disclose that mutations Glu478Gln and His539Phe in a complete HIV-1 RT/RNase H DNA fragment results in defective RNase activity without effecting RT activity.

Mizrahi, et al. (1990, Nucl. Acids. Res. 18: pp. 5359-5353) disclose additional mutations Asp443Asn and Asp498Asn in the RNase region of the pol gene which also results in defective RNase activity. The authors note that the Asp498Asn mutant was difficult to characterize due to instability of this mutant protein.

Leavitt, et al. (1993, *J. Biol. Chem.* 268: 2113-2119) disclose several mutations, including a Asp64Val mutation, which show differing effect on HIV-1 integrase (IN) activity.

Wiskerchen, et al. (1995, J. Virol. 69: 376-386) disclose singe and double mutants, including mutation of aspartic acid residues which effect HIV-1 IN and viral replication functions.

It would be of great import in the battle against AIDS to produce a prophylactic- and/or therapeutic-based HIV vaccine which generates a strong cellular immune response against an HIV infection. The present invention addresses and meets these needs by disclosing a class of adenovirus vaccines which, upon host administration, express codon optimized and modified versions of the HIV-1 genes, gag, pol and nef. These recombinant, replication-defective adenovirus vaccines may be administered to a host, such as a human, alone or as part of a combined modality regimen and/or prime-boost vaccination regimen with components of the present invention and/or a distinct viral HIV DNA vaccine, non-viral HIV DNA vaccine, HIV subunit vaccine, an HIV whole killed vaccine and/or a live attenuated HIV vaccine.

#### SUMMARY OF THE INVENTION

10

15

20

25

30

35

The present invention relates to enhanced replication-defective recombinant adenovirus vaccine vectors and associated recombinant, replication-deficient adenovirus vaccines which encode various forms of HIV-1 Gag, HIV-1 Pol, and/or HIV-1 Nef, including immunologically relevant modifications of HIV-1 Gag, HIV-1 Pol and HIV-1 Nef. The adenovirus vaccines of the present invention express HIV antigens and provide for improved cellular-mediated immune responses upon host administration. Potential vaccinees include but are not limited to primates and especially humans and non-human primates, and also include any non-human mammal of commercial or domestic veterinary importance. An effect of the improved recombinant adenovirus-based vaccines of the present invention should be a lower transmission rate to previously uninfected individuals (i.e., prophylactic applications) and/or reduction in the levels of the viral loads within an infected individual (i.e., therapeutic applications), so as to prolong the asymptomatic phase of HIV-1 infection. In particular, the present invention relates to adenoviral-based vaccines which encode various forms of codon optimized HIV-1 Gag (including but in no way limited to p55 versions of codon optimized full length (FL) Gag and tPA-Gag fusion proteins), HIV-1 Pol, HIV-1 Nef, and selected modifications of immunological relevance. The administration, intracellular delivery and expression of these adenovirus vaccines elicit a host CTL and Th response. The preferred replication-defective recombinant adenoviral vaccine vectors include but are not limited to synthetic DNA molecules which (1) encode codon optimized versions of wild type HIV-1 Gag; (2) encode codon optimized versions of HIV-1 Pol; (3) encode codon optimized versions of HIV-1 Pol fusion proteins; (4) encode codon optimized versions of modified HIV-1 Pol proteins and fusion proteins, including but not limited

to pol modifications involving residues within the catalytic regions responsible for RT, RNase and IN activity within the host cell; (5) encode codon optimized versions of wild type HIV-1 Nef; (6) codon optimized versions of HIV-1 Nef fusion proteins; and/or (7) codon optimized versions of HIV-1 Nef derivatives, including but not limited to nef modifications involving introduction of an amino-terminal leader sequence, removal of an amino-terminal myristylation site and/or introduction of dileucine motif mutations. The Nef-based fusion and modified proteins, disclosed within this specification and expressed from an adenoviral-based vector vaccine this specification, may possess altered trafficking and/or host cell function while retaining the ability to be properly presented to the host MHC I complex and in turn elicit a host CTL and Th response. Examples of HIV-1 Gag, Pol and/or Nef fusion proteins include but are not limited to fusion of a leader or signal peptide at the NH<sub>2</sub>-teriminal portion of the viral antigen coding region. Such a leader peptide includes but is not limited to a tPA leader peptide.

5

10

15

20

25

30

35

The adenoviral vector utilized in construction of the HIV-1 Gag-, HIV-1 Poland/or HIV-1 Nef- based vaccines of the present invention may comprise any replication-defective adenoviral vector which provides for enhanced genetic stability of the recombinant adenoviral genome through large scale production and purification of the recombinant virus. In other words, an HIV-1 Gag-, Pol- or Nef-based adenovirus vaccine of the present invention is a purified recombinant, replicationdefective adenovirus which is shown to be genetically stable through multiple passages in cell culture and remains so during large scale production and purification procedures. Such a recombinant adenovirus vector and harvested adenovirus vaccine lends itself to large scale dose filling and subsequent worldwide distribution procedures which will be demanded of an efficacious monovalent or multivalent HIV vaccine. The present invention meets this basic requirement with description of a replication-defective adenoviral vector and vectors derived therefrom, at least partially deleted in E1, comprising a wildtype adenovirus cis-acting packaging region from about base pair 1 to between from about base pair 342 (more preferably, 400) to about base pair 458 of the wildtype adenovirus genome. A preferred embodiment of the instant invention comprises base pairs 1-450 of a wildtype adenovirus. In other preferred embodiments, the replication -defective adenoviral vector has, in addition thereto, a region 3' to the E1-deleted region comprising base pairs 3511-3523. Basepairs 342-450 (more particularly, 400-450) constitute an extension of the 5'region of previously disclosed vectors carrying viral antigens, particularly HIV antigens (see, e.g., PCT International Application PCT/US00/18332, published

January 11, 2001 (WO 01/02067), which claims priority to U.S. Provisional Application Serial Nos. 60/142,631 and 60/148,981, filed 7/6/1999 and 8/13/1999, respectively; these documents herein incorporated by reference. Applicants have found that extending the 5' region further into the E1 gene into the disclosed vaccine vectors incorporated elements found to be important in optimizing the packaging of the virus.

As compared to previous vectors not comprising basepairs from about 1 to between from about base pair 342 (more preferably, 400) to about base pair 458 of the wildtype adenovirus genome, vectors comprising the above region exhibited enhanced growth characteristics, with approximately 5-10 fold greater amplification rates, a more potent virus effect, allowing lower doses of virus to be used to generate equivalent immunity; and a greater cellular-mediated immune response than replication-deficient vectors not comprising this region (basepairs 1-450). Even more important, adenoviral constructs derived therefrom are very stable genetically in large-scale production, particularly those comprising an expression cassette under the control of a hCMV promoter devoid of intron A. This is because Applicants have surprisingly found that the intron A portion of the hCMV promoter constituted a region of instability when employed in adenoviral vectors. Applicants have, therefore, identified an enhanced adenoviral vector which is particularly suited for use in gene therapy and nucleotide-based vaccine vectors which, favorably, lends itself to large scale propagation.

A preferred embodiment of this invention is a replication-defective adenoviral vector in accordance with the above description wherein the gene is inserted in the form of a gene expression cassette comprising (a) a nucleic acid encoding a protein or biologically active and/or immunologically relevant portion thereof; (b) a heterologous promoter operatively linked to the nucleic acid of part a); and, (c) a transcription terminator.

In preferred embodiments, the E1 gene, other than that contained within basepairs 1-450 or, alternatively, that contained within base pairs 1-450 and 3511-3523 has been deleted from the adenoviral vector, and the gene expression cassette has replaced the deleted E1 gene. In other preferred embodiments, the replication defective adenovirus genome does not have a functional E3 gene, or the E3 gene has been deleted. Most preferably, the E3 region is present within the adenoviral genome. Further preferred embodiments are wherein the gene expression cassette is in an E1 anti-parallel (transcribed in a 3' to 5' direction relative to the vector backbone)

orientation or, more preferably, an E1 parallel (transcribed in a 5' to 3' direction relative to the vector backbone) orientation.

5

10

15

20

25

30

35

Further embodiments relate to a shuttle plasmid vector comprising: an adenoviral portion and a plasmid portion, wherein said adenovirus portion comprises: a) a replication defective adenovirus genome, at least partially deleted in E1, comprising a wildtype adenovirus *cis*-acting packaging region from about base pair 1 to between from about base pair 342 (more preferably, 400) to about base pair 458 (preferably, 1-450) of the wildtype adenovirus genome and, preferably, in addition thereto, basepairs 3511-3523 of a wildtype adenovirus sequence; and b) a gene expression cassette comprising: (a) a nucleic acid encoding a protein or biologically active and/or immunologically relevant portion thereof; (b) a heterologous promoter operatively linked to the nucleic acid of part a); and (c) a transcription terminator and/or a polyadenylation site.

Other aspects of this invention include a host cell comprising said adenoviral vectors and/or said shuttle plasmid vectors; vaccine compositions comprising said vectors; and methods of producing the vectors comprising (a) introducing the adenoviral vector into a host cell which expresses adenoviral E1 protein, and (b) harvesting the resultant adenoviral vectors.

To this end, the present invention particularly relates to harvested recombinant, replication defective virus derived from a host cell, such as but not limited to 293 cells or PER.C6® cells, including but not limited to harvested virus related to any of the MRKAd5 vector backbones, with or without an accompanying transgene, including but not limited to the HIV-1 antigens described herein. An HIV-1 vaccine is represented by any harvested, recombinant adenovirus material which expresses any one or more of the HIV-1 antigens disclosed herein. This harvested material may then be purified, formulated and stored prior to host administration.

Another aspect of this invention is a method of generating a cellular immune response against a protein in an individual comprising administering to the individual an adenovirus vaccine vector comprising:

a) a recombinant, replication defective adenoviral vector, at least partially deleted in E1, comprising a wildtype adenovirus *cis*-acting adenovirus packaging region from about base pair 1 to between from about base pair 342 (more preferably, 400) to about base pair 458 (preferably, 1-450) and, preferably in addition thereto, base pairs 3511-3523 of a wildtype adenovirus sequence, and,

b) a gene expression cassette comprising:(i) a nucleic acid encoding a protein or biologically active and/or immunologically relevant portion thereof; (ii) a heterologous promoter operatively linked to the nucleic acid of part a); and (iii) a transcription terminator and/or a polyadenylation site.

5

10

15

20

25

30

35

In view of the efficacious nature of the adenoviral and/or DNA plasmid vaccines described herein, the present invention relates to all methodology regarding administration of one or more of these adenoviral and/or DNA plasmid vaccines to provide effective immunoprophylaxis, to prevent establishment of an HIV-1 infection following exposure to this virus, or as a post-HIV infection therapeutic vaccine to mitigate the acute HIV-1 infection so as to result in the establishment of a lower virus load with beneficial long term consequences. As discussed herein, such a treatment regimen may include a monovalent or multivalent composition, various combined modality applications, and/or a prime/boost regimen to as to optimize antigen expression and a concomitant cellular-mediated and/or humoral immune response upon inoculation into a living vertebrate tissue. Therefore, the present invention provides for methods of using the adenoviral and/or DNA plasmid vaccines disclosed herein within the various parameters disclosed herein as well as any additional parameters known in the art, which, upon introduction into mammalian tissue induces intracellular expression of the gag, pol and/or nef-based vaccines.

To this end, the present invention relates in part to methods of generating a cellular immune response in a vaccinee, preferably a human vaccinee, wherein the individual is given more than one administration of adenovirus vaccine vector, and it may be given in a regimen accompanied by the administration of a plasmid vaccine. The plasmid vaccine (also referred to herein as a "DNA plasmid vaccine" or "vaccine plasmid" comprises a nucleic acid encoding a protein or an immunologically relevant portion thereof, a heterologous promoter operably linked to the nucleic acid sequence, and a transcription terminator or a polyadenylation signal (such as bGH or SPA, respectively). There may be a predetermined minimum amount of time separating the administrations. The individual can be given a first dose of plasmid vaccine, and then a second dose of plasmid vaccine. Alternatively, the individual may be given a first dose of adenovirus vaccine, and then a second dose of adenovirus vaccine. In other embodiments, the plasmid vaccine is administered first, followed after a time by administration of the adenovirus vaccine. Conversely, the adenovirus vaccine may be administered first, followed by administration of plasmid vaccine after a time. In these embodiments, an individual may be given multiple doses of the same adenovirus serotype in either viral vector or plasmid form, or the virus may be of

differing serotypes. In the alternative, a viral antigen of interest can be first delivered via a viral vaccine other than an adenovirus-based vaccine, and then followed with the adenoviral vaccine disclosed. Alternative viral vaccines include but are not limited to pox virus and venezuelan equine encephilitis virus.

5

10

15

20

25

30

35

The present invention also relates to multivalent adenovirus vaccine compositions which comprise Gag, Pol and Nef components described herein; see, e.g., Example 29 and Table 25. Such compositions will provide for an enhanced cellular immune response subsequent to host administration, particularly given the genetic diversity of human MHCs and of circulating virus. Examples, but not limitations, include MRKAd5-vector based multivalent vaccine compositions which provide for a divalent (i.e., gag and nef, gag and pol, or pol and nef components) or a trivalent vaccine (i.e., gag, pol and nef components) composition. Such a mutlivalent vaccine may be filled for a single dose or may consist of multiple inoculations of each individually filled component; and may in addition be part of a prime/boost regimen with viral or non-viral vector vaccines as introduced in the previous paragraph. To this end, preferred compositions are MRKAd5 adenovirus used in combination with multiple, distinct HIV antigen classes. Each HIV antigen class is subject to sequence manipulation, thus providing for a multitude of potential vaccine combinations; and such combinations are within the scope of the present invention. The utilization of such combined modalities vaccine formulation and administration increase the probability of eliciting an even more potent cellular immune response when compared to inoculation with a single modality regimen.

The concept of a "combined modality" as disclosed herein also covers the alternative mode of administration whereby multiple HIV-1 viral antigens may be ligated into a proper shuttle plasmid for generation of a pre-adenoviral plasmid comprising multiple open reading frames. For example, a trivalent vector may comprise a gag-pol-nef fusion, in either a E3(-) or E3(+) background, preferably a E3 deleted backbone, or possibly a "2+1" divalent vaccine, such as a gag-pol fusion (i.e., codon optimized p55 gag and inactivated optimized pol; Example 29 and Table 25) within the same MRKAd5 backbone, with each open reading frame being operatively linked to a distinct promoter and transcription termination sequence. Alternatively, the two open reading frames may be operatively linked to a single promoter, with the open reading frames operatively linked by an internal ribosome entry sequence (IRES). Therefore, a multivalent vaccine delivered as a single, or possibly a second harvested recombinant, replication-deficient adenovirus is contemplated as part of the present invention.

Therefore, the adenoviral vaccines and plasmid DNA vaccines of this invention may be administered alone, or may be part of a prime and boost administration regimen. A mixed modality priming and booster inoculation scheme will result in an enhanced immune response, particularly if pre-existing anti-vector immune responses are present. This one aspect of this invention is a method of priming a subject with the plasmid vaccine by administering the plasmid vaccine at least one time, allowing a predetermined length of time to pass, and then boosting by administering the adenoviral vaccine. Multiple primings typically, 1-4, are usually employed, although more may be used. The length of time between priming and boost may typically vary from about four months to a year, but other time frames may be used. In experiments with rhesus monkeys, the animals were primed four times with plasmid vaccines, then were boosted 4 months later with the adenoviral vaccine. Their cellular immune response was notably higher than that of animals which had only received adenoviral vaccine. The use of a priming regimen may be particularly preferred in situations where a person has a pre-existing anti-adenovirus immune response.

10

15

20

25

30

35

It is an object of the present invention to provide for enhanced replication-defective recombinant adenoviral vaccine vector backbones. These recombinant adenoviral backbones may accept one or more transgenes, which may be passaged through cell culture for growth, amplification and harvest.

It is a further object to provide for enhanced replication-defective recombinant adenoviral vaccine vectors which encode various transgenes.

It is also an object of the present invention to provide for a harvested recombinant, replication-deficient adenovirus which shows enhanced growth and amplification rates while in combination with increased virus stability after continuous passage in cell culture. Such a recombinant adenovirus is particularly suited for use in gene therapy and nucleotide-based vaccine vectors which, favorably, lends itself to large scale propagation.

To this end, it is an object of the present invention to provide for (1) enhanced replication-defective recombinant adenoviral vaccine vectors as described herein which encode various forms of HIV-1 Gag, HIV-1 Pol, and/or HIV-1 Nef, including immunologically relevant modifications of HIV-1 Gag, HIV-1 Pol and HIV-1 Nef, and (2) harvested, purified recombinant replication-deficient adenovirus generated by passage of the adenoviral vectors of (1) through one or multiple passages through cell culture, including but not limited to passage through 293 cells or PER.C6® cells.

It is also an object of the present invention to provide for recombinant adenovirus harvested by one or multiple passages through cell culture. As relating to recombinant adenoviral vaccine vector, this recombinant virus is harvested and formulated for subsequent host administration.

5

10

15

20

25

30

35

It is also an object of the present invention to provide for replication-defective adenoviral vectors wherein at least one gene is inserted in the form of a gene expression cassette comprising (a) a nucleic acid encoding a protein or biologically active and/or immunologically relevant portion thereof; (b) a heterologous promoter operatively linked to the nucleic acid of part a); and, (c) a transcription terminator.

It is also an object of the present invention to provide for a host cell comprising said adenoviral vectors and/or said shuttle plasmid vectors; vaccine compositions comprising said vectors; and methods of producing the vectors comprising (a) introducing the adenoviral vector into a host cell which expresses adenoviral E1 protein, and (b) harvesting the resultant adenoviral vectors. It is a further object of the present invention to provide for methods of generating a cellular immune response against a protein in an individual comprising administering to the individual an adenovirus vaccine vector comprising a) a replication defective adenoviral vector, at least partially deleted in E1, comprising a wildtype adenovirus cis-acting packaging region from about base pair 1 to between from about base pair 342 (more preferably, 400) to about 450 (preferably, 1-450) and, preferably, 3511-3523 of a wildtype adenovirus sequence, and, b) a gene expression cassette comprising:(i) a nucleic acid encoding a protein or biologically active and/or immunologically relevant portion thereof; (ii) a heterologous promoter operatively linked to the nucleic acid of part a); and (iii) a transcription terminator and/or a polyadenylation site.

It is also an object of the present invention to provide various alternatives for vaccine administration regimes, namely administration of one or more adenoviral and/or DNA plasmid vaccines described herein to provide effective immunoprophylaxis for uninfected individuals or a therapeutic treatment for HIV infected patients. Such processes include but are not limited to multivalent HIV-1 vaccine compositions, various combined modality regimes as well as various prime/boost alternatives. These methods of administration, relating to vaccine composition and/or scheduled administration, will increase the probability of eliciting an even more potent cellular immune response when compared to inoculation with a single modality regimen.

As used throughout the specification and claims, the following definitions and abbreviations are used:

"HAART" refers to - highly active antiretroviral therapy -.

"first generation" vectors are characterized as being replication-defective.

They typically have a deleted or inactivated E1 gene region, and preferably have a deleted or inactivated E3 gene region as well.

"AEX" refers to Anion Exchange chromatography.

"QPA" refers to Quick PCR-based Potency Assay.

"bps" refers to basepairs.

5

10

15

20

25

30

35

"s" or "str" denotes that the transgene is in the E1 parallel or "straight" orientation.

"PBMCs" refers to peripheral blood monocyte cells.

"FL" refers to full length.

"FLgag" refers to a full-length optimized gag gene, as shown in Figure 2.

"Ad5-Flgag" refers to an adenovirus serotype 5 replication deficient virus which carries an expression cassette which comprises a full length optimized gag gene under the control of a CMV promoter.

"Promoter" means a recognition site on a DNA strand to which an RNA polymerase binds. The promoter forms an initiation complex with RNA polymerase to initiate and drive transcriptional activity. The complex can be modified by activating sequences such as enhancers or inhibiting sequences such as silencers.

"Leader" means a DNA sequence at the 5' end of a structural gene which is transcribed along with the gene. This usually results a protein having an N-terminal peptide extension, often referred to as a pro-sequences.

"Intron" means a section of DNA occurring in the middle of a gene which does not code for an amino acid in the gene product. The precursor RNA of the intron is excised and is therefore not transcribed into mRNA not translated into protein.

"Immunologically relevant" or "biologically active" means (1) with regards to a viral protein, that the protein is capable, upon administration, of eliciting a measurable immune response within an individual sufficient to retard the propagation and/or spread of the virus and/or to reduce the viral load present within the individual; or (2) with regards to a nucleotide sequence, that the sequence is capable of encoding for a protein capable of the above.

"Cassette" refers to a nucleic acid sequence which is to be expressed, along with its transcription and translational control sequences. By changing the cassette, a vector can express a different sequence.

"bGHpA" refers to the bovine growth hormone transcription terminator/polyadenylation sequence.

"tPAgag" refers to a fusion between the leader sequence of the tissue plasminogen activator leader sequence and an optimized HIV gag gene, as exemplified in Figure 30A-B, whether in a DNA or adenovirus-based vaccine vector.

Where utilized, "IA" or "inact" refers to an <u>inactivated</u> version of a gene (e.g. IApol).

"MCS" is "multiple cloning site".

5

10

15

20

25

30

35

In general, adenoviral constructs, gene constructs are named by reference to the genes contained therein. For example:

"Ad5 HIV-1 gag", also referred to as the original HIV-1 gag adenoviral vector, is a vector containing a transgene cassette composed of a hCMV intron A promoter, the full length version of the human codon-optimized HIV-1 gag gene, and the bovine growth hormone polyadenylation signal. The transgene was inserted in the E1 antiparallel orientation in an E1 and E3 deleted adenovector.

"MRK Ad5 HIV-1 gag" also referred to as "MRKAd5gag" or "Ad5gag2" is an adenoviral vector taught herein which is deleted of E1, comprises basepairs 1-450 and 3511-3523, and has a human codon-optimized HIV-1 gene in an E1 parallel orientation under the control of a CMV promoter without intron A. The construct also comprises a bovine growth hormone polyadenylation signal.

"pV1JnsHIVgag", also referred to as "HIVFLgagPR9901", is a plasmid comprising the CMV immediate-early (IE) promoter and intron A, a full-length codon-optimized HIV gag gene, a bovine growth hormone-derived polyadenylation and transcriptional termination sequence, and a minimal pUC backbone.

"pV1JnsCMV(no intron)-FLgag-bGHpA" is a plasmid derived from pV1JnsHIVgag which is deleted of the intron A portion of CMV and which comprises the full length HIV gag gene. This plasmid is also referred to as "pV1JnsHIVgag-bGHpA", pV1Jns-hCMV-FL-gag-bGHpA" and "pV1JnsCMV(no intron) + FLgag + bGHpA".

"pV1JnsCMV(no intron)-FLgag-SPA" is a plasmid of the same composition as pV1JnsCMV(no intron)-FLgag-bGHpA except that the SPA termination sequence replaces that of bGHpA. This plasmid is also referred to as "pV1Jns-HIVgag-SPA" and pV1Jns-hCMV-FLgag-SPA".

"pdelE1sp1A" is a universal shuttle vector with no expression cassette (i.e., no promoter or polyA). The vector comprises wildtype adenovirus serotype 5 (Ad5) sequences from bp 1 to bp 341 and bp 3524 to bp 5798, and has a multiple cloning

site between the Ad5 sequences ending 341 bp and beginning 3524 bp. This plasmid is also referred to as the original Ad 5 shuttle vector.

"MRKpdelE1sp1A" or "MRKpdelE1(Pac/pIX/pack450)" or

10

15

20

25

30

35

"MRKpdelE1(Pac/pIX/pack450)Cla1" is a universal shuttle vector with no expression cassette (i.e. no promoter or polyA) comprising wildtype adenovirus serotype 5 (Ad5) sequences from bp1 to bp450 and bp 3511 to bp 5798. The vector has a multiple cloning site between the Ad5 sequence ending 450 bp and beginning 3511 bp. This shuttle vector may be used to insert the CMV promoter and the bGHpA fragments in both the straight ("str". or E1 parallel) orientation or in the opposite (opp. or E1 antiparallel) orientation)

"MRKpdelE1(Pac/pIX/pack450)+CMVmin+BGHpA(str.)" is still another shuttle vector which is the modified vector that contains the CMV promoter (no intronA) and the bGHpA fragments. The expression unit containing the hCMV promoter (no intron A) and the bovine growth hormone polyadenylation signal has been inserted into the shuttle vector such that insertion of the gene of choice at a unique BgIII site will ensure the direction of transcription of the transgene will be Ad5 E1 parallel when inserted into the MRKpAd5(E1/E3+)Cla1 pre-plasmid. This shuttle vector, as shown in Figures 22 and 23, was used to insert the respective IApol and G2A,LLAA nef genes directly into.

"MRKpdelE1-CMV(no intron)-FLgag-bGHpA" is a shuttle comprising Ad5 sequences from basepairs 1-450 and 3511-5798, with an expression cassette containing human CMV without intron A, the full-length human codon-optimized HIV gag gene and bovine growth hormone polyadenylation signal. This plasmid is also referred to as "MRKpdelE1 shuttle +hCMV-FL-gag-BGHpA"

"MRKpAdHVE3+CMV(no intron)-FLgag-bGHpA" is an adenoviral vector comprising all Ad5 sequences except those nucleotides encompassing the E1 region (from 451-3510), a human CMV promoter without intron A, a full-length human codon-optimized HIV gag gene, and a bovine growth hormone polyadenylation signal. This vector is also referred to as "MRKpAdHVE3 + hCMV-FL-gag-BGHpA", "MRKpAd5HIV-1gag", "MRKpAd5gag", "pMRKAd5gag" or "pAd5gag2".

"pV1Jns-HIV-pol inact(opt)" or "pV1Jns-HIV IA pol (opt) is the inactivated Pol gene (contained within SEQ ID NO:3) cloned into the BgIII site of V1Jns (Figure 17A-C). As noted herein, various derivatives of HIV-1 pol may be cloned into a plasmid expression vector such as V1Jns or V1Jns-tPA, thus serving directly as DNA vaccine candidates or as a source for subcloning into an appropriate adenoviral vector.

"MRKpdel+hCMVmin+FL-pol+bGHpA(s)" is the "MRKpdelE1(Pac/pIX/pack450)+CMVmin+BGHpA(str.)" shuttle mentioned above which contains the IA pol gene is the proper orientation. This shuttle vector is used in a bacterial recombination with MRKpAd(E1-/E3+)Cla1.

"MRKpAd+hCMVmin+FL-pol+bGHpA(S)E3+", also referred to herein as "pMRKAd5pol", is the pre-adenovirus plasmid which comprises a CMV-pol inact(opt)-pGHpA construct. The construction of this pre-adenovirus plasmid is shown in Figure 22.

"pV1Jns/nef (G2A,LLAA)" or "V1Jns/opt nef (G2A,LLAA)" comprises codon optimized HTV-1 Nef wherein the open reading frame codes for modifications at the amino terminal myristylation site (Gly-2 to Ala-2) and substitution of the Leu-174-Leu-175 dileucine motif to Ala-174-Ala-175 (SEQ ID NO:13; which comprises an initiating methionine residue at nucleotides 12-14 and a "TAA" stop codon from nucleotides 660-662). This fragment is subcloned into the Bgl II site of V1Jns and/orV1Jns-tPA (Figures 16A-B). As noted above for HIV-1 pol, HIV-1 nef constructs may be cloned into a plasmid expression vector such as V1Jns or V1Jns-tPA, thus serving directly as DNA vaccine candidates or as a source for subcloning into an appropriate adenoviral vector.

"MRKpdelE1hCMVminFL-nefBGHpA(s)", also referred to herein as "pMRKAd5nef", is the pre-adenovirus plasmid which comprises a CMV-nef (G2A,LLAA) codon optimized sequence. The construction of this pre-adenovirus plasmid is shown in Figure 23.

# BRIEF DESCRIPTION OF THE FIGURES

5

10

15

20

25

30

35

Figure 1 shows the original HIV-1 gag adenovector (Ad5HIV-1gag). This vector is disclosed in PCT International Application No. PCT/US00/18332 (WO 01/02607) filed July 3, 2000, claiming priority to U.S. Provisional Application Serial No. 60/142,631, filed July 6, 1999 and U.S. Application Serial No. 60/148,981, filed August 13, 1999, all three applications which are hereby incorporated by reference.

Figure 2 shows the nucleic acid sequence (SEQ ID NO: 29) of the optimized human HIV-1 gag open reading frame.

Figure 3 shows diagrammatically the new transgene constructs in comparison with the original gag transgene.

Figure 4 shows the modifications made to the original adenovector backbone in the generation of the novel vectors of the instant invention.

Figure 5 shows the virus mixing experiments that were carried out to determine the effects of the addition made to the packaging signal region (Expt. #1) and the E3 gene on viral growth (Expt. #2). The bars denote the region of modifications made to the E1 deletion.

Figure 6 shows an autoradiograph of viral DNA analysis following the viral mixing experiments described in Examples 6 and 7.

5

10

15

20

25

30

35

Figures 7A, 7B and 7C are as follows: Figure 7A shows the hCMV-Flgag-bGHpA adenovectors constructed within the MRKpAdHVE3 and MRKpAdHVO adenovector backbones. Both E1 parallel and E1 antiparallel transgene orientation are represented. Figure 7B shows the hCMV-Flgag-SPA adenovectors constructed within the MRKpAdHVE3 and MRKpAdHVO adenovector backbones. Again, both E1 parallel and E1 antiparallel transgene orientation are represented. Figure 7C shows the mCMV-Flgag-bGHpA adenovectors constructed within the MRKpAdHVE3 and MRKpAdHVO adenovector backbones. Once again, both E1 parallel and E1 antiparallel transgene orientation are represented.

Figure 8A shows the experiment designed to test the effect of transgene orientation.

Figure 8B shows the experiments designed to test the effect of polyadenylation signal.

Figure 9 shows viral DNA from the four adenoviral vectors tested (Example 12) at P5, following *Bst*E11 digestion.

Figure 10 shows viral DNA analysis of passages 11 and 12 of MRKpAdHVE3, MRKAd5HIV-1gag, and MRKAd5HIV-1gagE3-.

Figure 11 shows viral DNA analysis (*Hind*III digestion) of passage 6 MRKpAdHVE3 and MRKAd5HIV-1gag used to initiate the viral competition study. The last two lanes are passage 11 analysis of duplicate passages of the competition study (each virus at MOI of 280 viral particles).

Figure 12 shows viral DNA analysis by *Hind* III digestion on high passage numbers for MRKAd5HIV-1gag in serum-containing media with collections made at specified times. The first lane shows the 1kb DNA size marker. The other lanes represent pre-plasmid control (digested with Pac1 and *Hind*III), MRKAd5HIV-1gag at P16, P19, and P21.

Figure 13 shows serum anti-p24 levels at 3 wks post i.m. immunization of balb/c mice (n=10) with varying doses of several Adgag constructs: (A) MRK Ad5 HIV-1 gag (through passage 5); (B) MRKAd5 hCMV-FLgag-bGHpA (E3-); (C) MRKAd5 hCMV-FLgag-SPA (E3+); (D) MRKAd5 mCMV-FLgag-bGHpA (E3+);

(E) research lot (293 cell-derived) of Ad5HIV-1 gag; and (F) clinical lot (Ad5gagFN0001) of Ad5HIV-1 gag. Reported are the geometric mean titers (GMT) for each cohort along with the standard error bars.

Figure 14 shows a restriction map of the pMRKAd5HIV-1gag vector.

Figures 15A-X illustrates the nucleotide sequence of the pMRKAd5HIV-1gag vector (SEQ ID NO:27.[coding] and SEQ ID NO:28 [non-coding]).

5

10

15

20

25

30

35

Figures 16A-B shows a schematic representation of DNA vaccine expression vectors V1Jns (A) and V1Jns-tPA (B), which are utilized for HIV-1 gag, pol and nef constructs in various DNA/viral vector combined modality regimens as disclosed herein.

Figures 17A-C shows the nucleotide (SEQ ID NO:3) and amino acid sequence (SEQ ID NO:4) of IA-Pol. Underlined codons and amino acids denote mutations, as listed in Table 1.

Figure 18 shows codon optimized nucleotide and amino acid sequences through the fusion junction of tPA-pol inact(opt) (contained within SEQ ID NOs: 7 and 8, respectively). The underlined portion represents the NH<sub>2</sub>-terminal region of IA-Pol.

Figures 19A-B show a nucleotide sequence comparison between wild type nef(jrfl) and codon optimized nef. The wild type nef gene from the jrfl isolate consists of 648 nucleotides capable of encoding a 216 amino acid polypeptide. WT, wild type sequence (SEQ ID NO:19); opt, codon-optimized sequence (contained within SEQ ID NO:1). The Nef amino acid sequence is shown in one-letter code (SEQ ID NO:2).

Figures 20A-C show nucleotide sequences at junctions between nef coding sequence and plasmid backbone of nef expression vectors V1Jns/nef (Figure 20A), V1Jns/nef(G2A,LLAA) (Figure 20B), V1Jns/tpanef (Figure 20C) and V1Jns/tpanef(LLAA) (Figure 20C, also). 5' and 3' flanking sequences of codon optimized nef or codon optimized nef mutant genes are indicated by bold/italic letters; nef and nef mutant coding sequences are indicated by plain letters. Also indicated (as underlined) are the restriction endonuclease sites involved in construction of respective nef expression vectors. V1Jns/tpanef and V1Jns/tpanef(LLAA) have identical sequences at the junctions.

Figure 21 shows a schematic presentation of nef and nef derivatives. Amino acid residues involved in Nef derivatives are presented. Glycine 2 and Leucine 174 and 175 are the sites involved in myristylation and dileucine motif, respectively. For both versions of the tpanef fusion genes, the putative leader peptide cleavage sites are

indicated with "\*", and a exogenous serine residue introduced during the construction of the mutants is underlined.

Figure 22 shows diagrammatically the construction of the pre-adenovirus plasmid construct, MRKAd5Pol.

Figure 23 shows diagrammatically the construction of the pre-adenovirus plasmid construct, MRKAd5Nef.

5

10

15

20

25

35

Figure 24 shows a comparison of clade B vs. clade C anti-gag T cell responses in clade B HIV-infected subjects.

Figure 25 shows a comparison of clade B vs. clade C anti-nef T cell responses in clade B HIV-infected subjects.

Figures 26A-AO illustrates the nucleotide sequence of the pMRKAd5HIV-1pol adenoviral vector (SEQ ID NO:32 [coding] and SEQ ID NO:33 [non-coding]), comprising the coding region of the inactivated pol gene (SEQ ID NO3).

Figures 27A-AM illustrates the nucleotide sequence of the pMRKAd5HIV-1 nef adenoviral vector (SEQ ID NO:34 [coding] and SEQ ID NO:35 [non-coding]), comprising the coding region of the inactivated pol gene (SEQ ID NO13).

Figure 28 shows the stability of MRKAd5 vectors comprising various promoter fragments (hCMV or mCMV) and terminations signals (bGH or SPA) in E3(+) or E3(-) backbones.

Figures 29A and B shows the anion-exchange HPLC viral particle concentrations of the freeze-thaw recovered cell associated virus at the 24, 36, 48, and 60 hpi time points (Figure 29A) and the timcourse QPA supernatant titers (Figure 29B) for MRKAd5gag, MRKAd5pol and MRKAd5nef.

Figure 30 shows the nucleotide sequence (SEQ ID NO:36) and amino acid sequence (SEQ ID NO:37) comprising the open reading frame of a representative tPA-gag fusion for use in the DNA and/or adenoviral vaccine disclosed herein.

Figure 31 shows the intracellular γIFN staining of PBMCs collected at week 10 (post DNA prime) and week 30 (post Ad boost). The cells were stimulated overnight in the presence or absence of the gag peptide pool. They were subsequently stained using fluorescence-tagged anti-CD3, anti-CD8, anti-CD4, and anti-γIFN monoclonal antibodies. Each plot shows all CD3+ T cells which were segregated in terms of positive staining for surface CD8 and γIFN production. The numbers in the upper right and lower right quadrants of each plot are the percentages of CD3<sup>+</sup> cells that were CD8<sup>+</sup>γIFN<sup>+</sup> and CD4<sup>+</sup>γIFN<sup>+</sup>, respectively.

Figure 32 shows a comparison of single-modality adenovirus immunization with DNA + adjuvant prime/adenovirus boost immunization.

Figures 33A-B show the nucleotide sequence (SEQ ID NO: 38) of the open reading frame for the gag-IApol fusion of Example 29.

Figures 34A-B show the protein sequence (SEQ ID NO:39) of the gag-IApol fustion frame.

5

10

15

20

25

30

35

# DETAILED DESCRIPTION OF THE INVENTION

A novel replication-defective, or "first generation," adenoviral vector suitable for use in gene therapy or nucleotide-based vaccine vectors is described. This vector is at least partially deleted in E1 and comprises a wildtype adenovirus cis-acting packaging region from about base pair 1 to between about base pair 342 (more preferably, 400) to about 458 (preferably, 1-450) and, preferably, 3511-3523 of a wild-type adenovirus sequence. It has been found that a vector of this description possesses enhanced growth characteristics, with approximately 5-10 fold greater amplification rates, and is more potent allowing lower doses of virus to be used to generate equivalent immunity. The vector, furthermore, generates a harvested recombinant adenovirus which shows greater cellular-mediated immune responses than replication-deficient vectors not comprising this region (basepairs 342-450). Adenoviral constructs derived from these vectors are, further, very stable genetically, particularly those comprising a transgene under the control of a hCMV promoter devoid of intron A. Viruses in accordance with this description were passaged continually and analyzed; see Example 12. Each virus analyzed maintained it correct genetic structure. Analysis was also carried out under propagation conditions similar to that performed in large scale production. Again, the vectors were found to possess enhanced genetic stability; see Figure 12. Following 21 passages, the viral DNA showed no evidence of rearrangement, and was highly reproducible from one production lot to the next. The outcome of all relevant tests indicate that the adenoviral vector is extremely well suited for large-scale production of recombinant, replication-deficient adenovirus, as shown herein with the data associated with Figure 28.

A preferred adenoviral vector in accordance with this description is a vector comprising basepairs 1-450, which is deleted in E3. This vector can accommodate up to approximately 7,500 base pairs of foreign DNA inserts (or exogenous genetic material). Another preferred vector is one retaining E3 which comprises basepairs 1-450. A preferred vector of this description is an E3+ vector comprising basepairs 1-450 and 3511-3523. This vector, when deleted of the region spanning basepairs 451-3510, can accommodate up to approximately, 4,850 base pairs of foreign DNA inserts

(or exogenous genetic material). The cloning capacities of the above vectors have been determined using 105% of the wildtype Ad5 sequence as the upper genome size limit.

Wildtype adenovirus serotype 5 is used as the basis for the specific basepair numbers provided throughout the specification. The wildtype adenovirus serotype 5 sequence is known and described in the art; see, Chroboczek et al., 1992 J. Virology 186:280, which is hereby incorporated by reference. Accordingly, a particular embodiment of the instant invention is a vector based on the adenovirus serotype 5 sequence. One of skill in the art can readily identify the above regions in other adenovirus serotypes (e.g., serotypes 2, 4, 6, 12, 16, 17, 24, 31, 33, and 42), regions defined by basepairs corresponding to the above basepair positions given for adenovirus serotype 5. Accordingly, the instant invention encompasses all adenoviral vectors partially deleted in E1 comprising basepairs corresponding to 1-450 (particularly, 342-450) and, preferably, 3511-3523 of a wild-type adenovirus serotype 5 (Ad5) nucleic acid sequence. Particularly preferred embodiments of the instant invention are those derived from adenoviruses like Ad5 which are classified in subgroup C (e.g., Ad2).

Vectors in accordance with the instant invention are at least partially deleted in E1. Preferably the E1 region is completely deleted or inactivated. Most preferably, the region deleted of E1 is within basepairs 451-3510. It is to be noted that the extended 5' and 3' regions of the disclosed vectors are believed to effectively reduce the size of the E1 deletion of previous constructs without overlapping any part of the E1A/E1B gene present in the cell line used, i.e., the PER.C6® cell line transfected with base pairs 459-3510. Overlap of adenoviral sequences is avoided because of the possibility of recombination. One of ordinary skill in the art can certainly appreciate that the instant invention can, therefore, be modified if a different cell line transfected with a different segment of adenovirus DNA is utilized. For purposes of exemplification, a 5' region of base pairs 1 to up to 449 is more appropriate if a cell line is transfected with adenoviral sequence from base pairs 450-3510. This holds true as well in the consideration of segments 3' to the E1 deletion.

Preferred embodiments of the instant invention possess an intact E3 region (i.e., an E3 gene capable of encoding a functional E3). Alternate embodiments have a partially deleted E3, an inactivated E3 region, or a sequence completely deleted of E3. Applicants have found, in accordance with the instant invention, that virus comprising the E3 gene were able to amplify more rapidly compared with virus not comprising an E3 gene; see Figure 6 wherein a diagnostic CsCl band corresponding to the E3+ virus

tested (5,665 bp) was present in greater amount compared with the diagnostic band of 3,010 bp corresponding to the E3- virus. These results were obtained following a virus competition study involving mixing equal MOI ratio (1:1) of adenovectors both comprising the E3 gene and not comprising the E3 gene. This increased amplification capacity of the E3+ adenovectors was subsequently confirmed with growth studies; see Table 4A, wherein the E3+ virus exhibit amplification ratios of 470, 420 and 320 as compared with the 115 and 40-50 of the E3- constructs.

As stated above, vectors in accordance with the instant invention can accommodate up to approximately 4,850 base pairs of exogenous genetic material for an E3+ vector and approximately 7,500 base pairs for an E3- vector. Preferably, the insert brings the adenoviral vector as close as possible to a wild-type genomic size (e.g., for Ad5, 35,935 basepairs). It is well known that adenovirus amplifies best when they are close to their wild-type genomic size.

10

15

20

25

30

35

The genetic material can be inserted in an E1-parallel or an E1 anti-parallel orientation, as such is illustrated in Figure 7A, 7B, 7C and Figure 8A. Particularly preferred embodiments of the instant invention, have the insert in an E1-parallel orientation. Applicants have found, via competition experiments with plasmids containing transgenes in differing orientation (Figure 8A), that vector constructs with the foreign DNA insert in an E1-parallel orientation amplify better and actually outcompete E1-antiparallel-oriented transgenes. Viral DNA analysis of the mixtures at passage 3 and certainly at passage 6, showed a greater ratio of the virus carrying the transgene in the E1 parallel orientation as compared with the E1 anti-parallel version. By passage 10, the only viral species observed was the adenovector with the transgene in the E1 parallel orientation for both transgenes tested.

Adenoviral vectors in accordance with the instant invention are particularly well suited to effectuate expression of desired proteins, one example of which is an HIV protein, particularly an HIV full length gag protein. Exogenous genetic material encoding a protein of interest can exist in the form of an expression cassette. A gene expression cassette preferably comprises (a) a nucleic acid encoding a protein of interest; (b) a heterologous promoter operatively linked to the nucleic acid encoding the protein; and (c) a transcription terminator.

The transcriptional promoter is preferably recognized by an eukaryotic RNA polymerase. In a preferred embodiment, the promoter is a "strong" or "efficient" promoter. An example of a strong promoter is the immediate early human cytomegalovirus promoter (Chapman et al, 1991 *Nucl. Acids Res*19:3979-3986, which is incorporated by reference), preferably without intronic sequences. Most preferred

for use within the instant adenoviral vector is a human CMV promoter without intronic sequences, like intron A. Applicants have found that intron A, a portion of the human cytomegalovirus promoter (hCMV), constitutes a region of instability for adenoviral vectors. CMV without intron A has been found to effectuate (Examples 1-3) comparable expression capabilities in vitro when driving HIV gag expression and, furthermore, behaved equivalently to intron A-containing constructs in Balb/c mice in vivo with respect to their antibody and T-cell responses at both dosages of plasmid DNA tested (20 µg and 200 µg). Those skilled in the art will appreciate that any of a number of other known promoters, such as the strong immunoglobulin, or other eukaryotic gene promoters may also be used, including the EF1 alpha promoter, the murine CMV promoter, Rous sarcoma virus (RSV) promoter, SV40 early/late promoters and the beta-actin promoter.

5

10

15

20

25

30

35

In preferred embodiments, the promoter may also comprise a regulatable sequence such as the Tet operator sequence. This would be extremely useful, for example, in cases where the gene products are effecting a result other than that desired and repression is sought.

The combination of the CMV promoter (devoid of the intron A region) with the BGH terminator is particularly preferred although other promoter/terminator combinations in the context of FG adenovirus may also be used.

Other embodiments incorporate a leader or signal peptide into the transgene. A preferred leader is that from the tissue-specific plasminogen activator protein, tPA. Examples include but are not limited to the various tPA-gag, tPA-pol and tPA-nef adenovirus-based vaccines disclosed throughout this specification.

In view of the improved adenovirus vectors described herein, an essential portion of the present invention are adenoviral-based HIV vaccines comprising said adenovirus backbones which may be administered to a mammalian host, preferably a human host, in either a prophylactic or therapeutic setting. The HIV vaccines of the present invention, whether administered alone or in combination regimens with other viral- or non-viral-based DNA vaccines, should elicit potent and broad cellular immune responses against HIV that will either lessen the likelihood of persistent virus infection and/or lead to the establishment of a clinically significant lowered virus load

subject to HIV infection or in combination with HAART therapy, mitigate the effects of previously established HIV infection (antiviral immunotherapy(ARI)). While any HIV antigen (e.g., gag, pol, nef, gp160, gp41, gp120, tat, rev, etc.) may be utilized in the herein described recombinant adenoviral vectors, preferred embodiments include the codon optimized p55 gag antigen (herein exemplified as MRKAd5gag), pol and nef. Sequences based on different Clades of HIV-1 are suitable for use in the instant invention, most preferred of which are Clade B and Clade C. Particularly preferred embodiments are those sequences (especially, codon-optimized sequences) based on concensus Clade B sequences. Preferred versions of the MRKAd5pol and MRKAd5nef series of adenoviral vaccines will encode modified versions of pol or nef, as discussed herein. Preferred embodiments of the MRKAd5HIV-1 vectors carrying HIV envelope genes and modifications thereof comprise the HIV codon-optimized env sequences of PCT International Applications PCT/US97/02294 and PCT/US97/10517, published August 28, 1997 (WO 97/31115) and December 24, 1997, respectively; both documents of which are hereby incorporated by reference.

5

10

15

20

25

30

35

A most preferred aspect of the instant invention is the disclosed use of the adenoviral vector described above to effectuate expression of HIV gag. Sequences for many genes of many HIV strains are publicly available in GENBANK and primary, field isolates of HIV are available from the National Institute of Allergy and Infectious Diseases (NIAID) which has contracted with Quality Biological (Gaithersburg, MD) to make these strains available. Strains are also available from the World Health Organization (WHO), Geneva Switzerland. It is preferred that the gag gene be from an HIV-1 strain (CAM-1; Myers et al, eds. "Human Retroviruses and AIDS: 1995, IIA3-IIA19, which is hereby incorporated by reference). This gene closely resembles the consensus amino acid sequence for the clade B (North American/European) sequence. Therefore, it is within the purview of the skilled artisan to choose an appropriate nucleotide sequence which encodes a specific HIV gag antigen, or immunologically relevant portion thereof. As shown in Example 25, a clade B or clade C based p55 gag antigen will potentially be useful on a global scale. As noted herein, the transgene of choice for insertion in to a DNA or MRKAd-based adenoviral vector of the present invention is a codon optimized version of p55 gag. Such a MRKAd5gag adenoviral vector is documented in Example 11 and is at least referred to herein as MRKAd5HIV-1gag. Of course, additional versions are contemplated, including but not limited to modifications such as promoter (e.g., mCMV for hCMV) and/or pA-terminations signal (SPA for bGH) switching, as well as generating MRK Ad5 backbones with or without deletion of the Ad5 E3 gene.

The present invention also relates a series of MRKAd5pol-based adenoviral vaccines which are shown herein to generate cellular immune responses subsequent to administration in mice and non-human primate studies. Several of the MRKAd5pol series are exemplified herein. One such adenoviral vector is referred to as MRKAd5hCMV-inact opt pol(E3+), which comprises the MRKAd5 backbone, the hCMV promoter (no intron A), an inactivated pol transgene, and contains the Ad5 E3 gene in the adenoviral backbone. A second exemplified pre-adenovirus plasmid and concomitant virus is referred to as MRKAd5hCMV-inact opt pol(E3-), which is identical to the former adenoviral vector except that the E3 is deleted. Both constructions contain a codon optimized, inactivated version of HIV-1 Pol, wherein at least the entire coding region is disclosed herein as SEQ ID NO:3 and the expressed protein is shown as SEQ ID NO:4 (see also Figure 17A-C and Table 1, which show targeted deletion for inactivated pol. This and other preferred codon optimized versions of HIV Pol as disclosed herein are essentially as described in U.S. Application Serial No. 09/745,221, filed December 21, 2000 and PCT International Application PCT/US00/34724, also filed December 21, 2000, both documents which are hereby incorporated by reference. As disclosed in the above-mentioned documents, the open reading frame for these codon-optimized HIV-1 Pol-based DNA vaccines are represented by codon optimized DNA molecules encoding codon optimized HIV-1 Pol (e.g. SEQ ID NO:2), codon optimized HIV-1 Pol fused to an amino terminal localized leader sequence (e.g. SEQ ID NO:6), and especially preferable, and exemplified by the MRKAd5-Pol construct in e.g., Example 19, biologically inactivated pol ("inact opt Pol"; e.g., SEQ ID NO:4) which is devoid of significant PR, RT, RNase or IN activity associated with wild type Pol. In addition, a construct related to SEO ID NO:4 is contemplated which contains a leader peptide at the amino terminal region of the IA Pol protein. A specific construct is ligated within an appropriate DNA plasmid vector containing regulatory regions operatively linked to the respective HIV-1 Pol coding region, with or without a nucleotide sequence encoding a functional leader peptide. To this end, various HIV-1 Pol constructs disclosed herein relate to open reading frames for cloning to the enhanced first generation Ad vectors of the present invention (such a series of MRKAd5pol adenoviral vaccine vectors), including but not limited to wild type Pol (comprising the DNA molecule encoding WT opt Pol, as set forth in SEQ ID NO:2), tPA-opt WTPol, (comprising the DNA molecule encoding tPA Pol, as set forth in SEQ ID NO:6), inact opt Pol (comprising the DNA molecule encoding IA Pol, as set forth in SEQ ID NO:4), and tPA-inact opt Pol, (comprising the DNA molecule encoding tPA-inact opt

10

15

20

25

35

Pol, as set forth in SEQ ID NO:8). The pol-based versions of enhanced first generation adenovirus vaccines elicit CTL and Th cellular immune responses upon administration to the host, including primates and especially humans. As noted in the above, an effect of the cellular immune-directed vaccines of the present invention should be a lower transmission rate to previously uninfected individuals and/or reduction in the levels of the viral loads within an infected individual, so as to prolong the asymptomatic phase of HIV-1 infection.

10

15

20

25

30

35

The present invention further relates to a series of MRKAd5nef-based adenoviral vaccines which, similar to HIV gag and pol antigens, generate cellular immune responses subsequent to administration in mice and non-human primate studies. The MRKAd5nef series are exemplified herein by utilizing the improved MRK adenoviral backbone in combination with modified versions of HIV nef. These exemplified MRKAd5nef vectors are as follows: (1) MRKAd5hCMVnef(G2A,LLAA) (E3+), which comprises the improved MRKAd5 backbone, a human CMV promoter an intact Ad5 E3 gene and a modified nef gene: (2) MRKAd5mCMVnef(G2A,LLAA) (E3+), which is the same as (1) above but substituting a murine CMV promoter for a human CMV promoter; and (3) MRKAd5mCMV-tpanef(LLAA) (E3+), which is the same as (2) except that the nef transgene is tpanef(LLAA). Codon optimized versions of HIV-1 Nef and HIV-1 Nef modifications are essentially as described in U.S. Application Serial No. 09/738,782, filed December 15, 2000 and PCT International Application PCT/US00/34162, also filed December 15, 2000, both documents which are hereby incorporated by reference. Particular embodiments of codon optimized Nef and Nef modifications relate to a DNA molecule encoding HIV-1 Nef from the HIV-1 jfrl isolate wherein the codons are optimized for expression in a mammalian system such as a human. The DNA molecule which encodes this protein is disclosed herein as SEQ ID NO:9, while the expressed open reading frame is disclosed herein as SEQ ID NO:10. Another embodiment of Nef-based coding regions for use in the adenoviral vectors of the present invention comprise a codon optimized DNA molecule encoding a protein containing the human plasminogen activator (tpa) leader peptide fused with the NH2-terminus of the HIV-1 Nef polypeptide. The DNA molecule which encodes this protein is disclosed herein as SEO ID NO:11, while the expressed open reading frame is disclosed herein as SEQ ID NO:12. Another modified Nef optimized coding region relates to a DNA molecule encoding optimized HIV-1 Nef wherein the open reading frame codes for modifications at the amino terminal myristylation site (Gly-2 to Ala-2) and substitution of the Leu-174-Leu-175 dileucine motif to Ala-174-Ala-175, herein

described as opt nef (G2A, LLAA). The DNA molecule which encodes this protein is disclosed herein as SEQ ID NO:13, while the expressed open reading frame is disclosed herein as SEQ ID NO:14. MRKAd5nef vectors (1) MRKAd5hCMV-nef(G2A,LLAA) (E3+) and (2) MRKAd5mCMV-nef(G2A,LLAA) (E3+) contain this transgene. An additional embodiment relates to a DNA molecule encoding optimized HIV-1 Nef wherein the amino terminal myristylation site and dileucine motif have been deleted, as well as comprising a tPA leader peptide. This DNA molecule, opt tpanef (LLAA), comprises an open reading frame which encodes a Nef protein containing a tPA leader sequence fused to amino acid residue 6-216 of HIV-1 Nef (jfrl), wherein Leu-174 and Leu-175 are substituted with Ala-174 and Ala-175, herein referred to as opt tpanef (LLAA) is disclosed herein as SEQ ID NO:15, while the expressed open reading frame is disclosed herein as SEQ ID NO:16. The MRKAd5nef vector "MRKAd5mCMV-tpanef(LLAA) (E3+)" contains this transgene.

10

15

20

25

30

35

Along with the improved MRKAd5gag adenovirus vaccine vector described herein, generation of a MRKAd5pol and MRKAd5nef adenovirus vector provide for enhanced HIV vaccine capabilities. Namely, the generation of this trio of adenoviral vaccine vectors, all shown to generate effective cellular immune responses subsequent to host administration, provide for the ability to administer these vaccine candidates not only alone, but preferably as part of a divalent (i.e., gag and nef, gag and pol, or pol and nef components) or a trivalent vaccine (i.e., gag, pol and nef components). Therefore, a preferred aspect of the present invention are vaccine formulations and associated methods of administration and concomitant generation of host cellular immune responses associated with formulating three separate series of MRKAd5based adenoviral vector vaccines. Of course, this MRKAd5 vaccine series based on distinct HIV antigens promotes expanded opportunities for formulation of a divalent or trivalent vaccine, or possibly administration of separate formulations of one or more monovalent or divalent formulations within a reasonable window of time. It is also within the scope of the present invention to embark on combined modality regimes which include multiple but distinct components from a specific antigen. An example, but certainly not a limitation, would be separate MRKAd5pol vectors, with one vaccine vector expressing wild type Pol (SEQ ID NO:2) and another MRKAd5pol vector expressing inactivated Pol (SEQ ID NO:6). Another example might be separate MRKAd5nef vectors, with one vaccine vector expressing the tPA/LLAA version of Nef (SEQ ID NO:16) and another MRKAd5nef vector expressing the G2A,LLAA modified version of Nef (SEQ ID NO:14). Therefore, the MRKAd5 adenoviral vectors of the present invention may be used in combination

with multiple, distinct HIV antigen classes. Each HIV antigen class is subject to sequence manipulation, thus providing for a multitude of potential vaccine combinations; and such combinations are within the scope of the present invention. The utilization of such combined modalities vaccine formulation and administration increase the probability of eliciting an even more potent cellular immune response when compared to inoculation with a single modality regimen.

5

10

15

20

25

30

35

The present invention also relates to application of a mono-, dual-, or tri-modality administration regime of the MRKAd5gag, pol and nef adenoviral vaccine series in a prime/boost vaccination schedule. This prime/boost schedule may include any reasonable combination of the MRKAd5gag, pol and nef adenoviral vaccine series disclosed herein. In addition, a prime/boost regime may also involve other viral and/or non-viral DNA vaccines. A preferable addition to an adenoviral vaccine vector regime includes but is not limited to plasmid DNA vaccines, especially DNA plasmid vaccines that contain at least one of the codon optimized gag, pol and nef constructions, as disclosed herein.

Therefore, one aspect of this invention is the administration of the adenoviral vector containing the optimized gag gene in a prime/boost regiment in conjunction with a plasmid DNA encoding gag. To distinguish this plasmid from the adenoviralcontaining shuttle plasmids used in the construction of an adenovirus vector, this plasmid will be referred to as a "vaccine plasmid" or "DNA plasmid vaccine". Preferred vaccine plasmids for use in this administration protocol are disclosed in pending U.S. patent application 09/017,981, filed February 3, 1998 and WO98/34640, published August 13, 1998, both of which are hereby incorporated by reference. Briefly, the preferred vaccine plasmid is designated V1Jns-FLgag, which expresses the same codon-optimized gag gene as the adenoviral vectors of this invention (see Figure 2 for the nucleotide sequence of the exemplified optimized codon version of full length p55 gag). The vaccine plasmid backbone, designated V1Jns contains the CMV immediate-early (IE) promoter and intron A, a bovine growth hormone-derived polyadenylation and transcription termination sequence as the gene expression regulatory elements, and a minimal pUC backbone; see Montgomery et al., 1993, DNA Cell Biol. 12:777-783. The pUC sequence permits high levels of plasmid production in E. coli and has a neomycin resistance gene in place of an ampicillin resistance gene to provide selected growth in the presence of kanamycin. Alternatively, a vaccine plasmid which has the CMV promoter deleted of intron A can be used. Those of skill in the art will recognize that alternative vaccine plasmid

vectors may be easily substituted for these specific constructs, and this invention specifically envisions use of such alternative plasmid DNA vaccine vectors.

10

15

20

25

30

35

Another aspect of the present invention is a prime/boost regimen which includes a vaccine plasmid which encodes an HIV pol antigen, preferably a codon optimized form of pol and also preferably a vaccine plasmid which comprises a nucleotide sequence which encodes a Pol antigen selected from the group of Pol antigens as shown in SEQ ID NOs: 2, 4, 6 and 8. The variety of potential DNA plasmid vaccines which encode various biologically active forms of HIV-1 Pol, wherein administration, intracellular delivery and expression of the HIV-1 Pol gene of interest elicits a host CTL and Th response. The preferred synthetic DNA molecules of the present invention encode codon optimized wild type Pol (without Pro activity) and various codon optimized inactivated HIV-1 Pol proteins. The HIV-1 pol open reading disclosed herein are especially preferred for pharmaceutical uses, especially for human administration as delivered via a recombinant adenoviral vaccine, especially an enhanced first generation recombinant adenoviral vaccine as described herein. Several embodiments of this portion of the invention are provided in detail below, namely DNA molecules which comprise a HIV-1 pol open reading frame, whether encoding full length pol or a modification or fusion as described herein, wherein the codon usage has been optimized for expression in a mammal, especially a human. Again, these DNA sequences are positioned appropriately within a recombinant adenoviral vector, such as the exemplified recombinant adenoviral vector described herein, so as to promote expression of the respective HIV-1 Pol gene of interest, and subsequent to administration, elicit a host CTL and Th response. Again, these preferred, but in no way limiting, pol genes are as disclosed herein and essentially as described in U.S. Application Serial No. 09/745,221, filed December 21, 2000 and PCT International Application PCT/US00/34724, also filed December 21, 2000, both documents which are hereby incorporated by reference.

A third series of vaccine plasmids which are useful in a combined modality and/or prime/boost regimen are vaccine plasmids which encode an HIV nef antigen or biologically and/or immunologically relevant modification thereof. As noted elsewhere, preferred vaccine plasmids contain a codon optimized form of nef and also preferably comprise a nucleotide sequence which encodes a Nef antigen selected from the group of Nef antigens as shown in SEQ ID NOs: 10, 12, 14 and 16. These preferred nef coding regions are disclosed herein, as well as being described in U.S. Application Serial No. 09/738,782, filed December 15, 2000 and PCT International

Application PCT/US00/34162, also filed December 15, 2000, both documents which are hereby incorporated by reference.

5

10

15

20

25

30

35

Therefore, the adenoviral vaccines and plasmid DNA vaccines of this invention may be administered alone, or may be part of a prime and boost administration regimen. A mixed modality priming and booster inoculation scheme will result in an enhanced immune response, particularly is pre-existing anti-vector immune responses are present. This one aspect of this invention is a method of priming a subject with the plasmid vaccine by administering the plasmid vaccine at least one time, allowing a predetermined length of time to pass, and then boosting by administering the adenoviral vaccine. Multiple primings typically, 1-4, are usually employed, although more may be used. The length of time between priming and boost may typically vary from about four months to a year, but other time frames may be used. In experiments with rhesus monkeys, the animals were primed four times with plasmid vaccines, then were boosted 4 months later with the adenoviral vaccine. Their cellular immune response was notably higher than that of animals which had only received adenoviral vaccine. The use of a priming regimen may be particularly preferred in situations where a person has a pre-existing anti-adenovirus immune response.

Furthermore and in the alternative, multiple HTV-1 viral antigens, such as the MRKAd5 adenoviral vaccines disclosed herein, may be ligated into a proper shuttle plasmid for generation of a pre-adenoviral plasmid comprising multiple open reading frames. For example a trivalent vector may comprise a gag-pol-nef fusion, in either a E3(-) or E3(+) background, preferably a E3 deleted backbone, or possible a "2+1" divalent vaccine, such as a gag-pol fusion (i.e., codon optimized p55 gag and inactivated optimized pol; Example 29 and Table 25) within the same MRKAd5 backbone, with each open reading frame being operatively linked to a distinct promoter and transcription termination sequence. Alternatively, the two open reading frames may be operatively linked to a single promoter, with the open reading frames operatively linked by an internal ribosome entry sequence (IRES), as disclosed in International Publication No. WO 95/24485, which is hereby incorporated by reference. Figure 9 shows that the use of multiple promoters and termination sequences provide for similar growth properties, while Figure 28 shows that these MRKAd5gag-based vectors are also stable at least through passage 21. In the absence of the use of IRES-based technology, it is preferred that a distinct promoter be used to support each respective open reading frame, so as to best preserve vector stability. As examples, and certainly not as limitations, potential multiple transgene vaccines may

include a three transgene vector such as hCMV-gagpol-bGHpA + mCMV-nef-SPA in an E3 deleted backbone or hCMV-gagpol-bGHpA + mCMV-nef-SPA(E3+). Potential "2+1" divalent vaccines of the present invention might be a hCMV-gagbGHpA + mCMV-nef-SPA in an E3+ backbone (vector #1) in combination with hCMV-pol-bGHpA in an E3+ backbone (vector #2), with all transgenes in the E1 parallel orientation. Fusion constructs other than the gag-pol fusion described above are also suitable for use in various divalent vaccine strategies and can be composed of any two HIV antigens fused to one another (e.g.,, nef-pol and gag-nef). These adenoviral compositions are, as above, preferably delivered along with an adenoviral composition comprising an additional HIV antigen in order to diversify the immune response generated upon administration. Therefore, a multivalent vaccine delivered in a single, or possible second, adenoviral vector is certainly contemplated as part of the present invention. Again, this mode of administration is another example of whereby an efficaceous adenovirus-based HIV-1 vaccine may be administered via a combined modality regime. It is important to note, however, that in terms of deciding on an insert for the disclosed adenoviral vectors, due consideration must be dedicated to the effective packaging limitations of the adenovirus vehicle. Adenovirus has been shown to exhibit an upper cloning capacity limit of approximately 105% of the wildtype Ad5 sequence.

10

15

20

25

30

35

Regardless of the gene chosen for expression, it is preferred that the sequence be "optimized" for expression in a human cellular environment. A "triplet" codon of four possible nucleotide bases can exist in 64 variant forms. That these forms provide the message for only 20 different amino acids (as well as transcription initiation and termination) means that some amino acids can be coded for by more than one codon. Indeed, some amino acids have as many as six "redundant", alternative codons while some others have a single, required codon. For reasons not completely understood, alternative codons are not at all uniformly present in the endogenous DNA of differing types of cells and there appears to exist variable natural hierarchy or "preference" for certain codons in certain types of cells. As one example, the amino acid leucine is specified by any of six DNA codons including CTA, CTC, CTG, CTT, TTA, and TTG (which correspond, respectively, to the mRNA codons, CUA, CUC, CUG, CUU, UUA and UUG). Exhaustive analysis of genome codon frequencies for microorganisms has revealed endogenous DNA of E. coli most commonly contains the CTG leucine-specifying codon, while the DNA of yeasts and slime molds most commonly includes a TTA leucine-specifying codon. In view of this hierarchy, it is generally held that the likelihood of obtaining high levels of expression of a leucine-

rich polypeptide by an *E. coli* host will depend to some extent on the frequency of codon use. For example, a gene rich in TTA codons will in all probability be poorly expressed in *E. coli*, whereas a CTG rich gene will probably highly express the polypeptide. Similarly, when yeast cells are the projected transformation host cells for expression of a leucine-rich polypeptide, a preferred codon for use in an inserted DNA would be TTA.

5

10

15

20

25

30

35

The implications of codon preference phenomena on recombinant DNA techniques are manifest, and the phenomenon may serve to explain many prior failures to achieve high expression levels of exogenous genes in successfully transformed host organisms—a less "preferred" codon may be repeatedly present in the inserted gene and the host cell machinery for expression may not operate as efficiently. This phenomenon suggests that synthetic genes which have been designed to include a projected host cell's preferred codons provide a preferred form of foreign genetic material for practice of recombinant DNA techniques. Thus, one aspect of this invention is an adenovirus vector or adenovirus vector in some combination with a vaccine plasmid where both specifically include a gene which is codon optimized for expression in a human cellular environment. As noted herein, a preferred gene for use in the instant invention is a codon-optimized HIV gene and, particularly, HIV gag, pol or nef.

Adenoviral vectors in accordance with the instant invention can be constructed using known techniques, such as those reviewed in Hitt et al, 1997 "Human Adenovirus Vectors for Gene Transfer into Mammalian Cells" Advances in Pharmacology 40:137-206, which is hereby incorporated by reference.

In constructing the adenoviral vectors of this invention, it is often convenient to insert them into a plasmid or shuttle vector. These techniques are known and described in Hitt et al., *supra*. This invention specifically includes both the adenovirus and the adenovirus when inserted into a shuttle plasmid.

Preferred shuttle vectors contain an adenoviral portion and a plasmid portion. The adenoviral portion is essentially the same as the adenovirus vector discussed supra, containing adenoviral sequences (with non-functional or deleted E1 and E3 regions) and the gene expression cassette, flanked by convenient restriction sites. The plasmid portion of the shuttle vector often contains an antibiotic resistance marker under transcriptional control of a prokaryotic promoter so that expression of the antibiotic does not occur in eukaryotic cells. Ampicillin resistance genes, neomycin resistance genes and other pharmaceutically acceptable antibiotic resistance markers may be used. To aid in the high level production of the polynucleotide by

fermentation in prokaryotic organisms, it is advantageous for the shuttle vector to contain a prokaryotic origin of replication and be of high copy number. A number of commercially available prokaryotic cloning vectors provide these benefits. It is desirable to remove non-essential DNA sequences. It is also desirable that the vectors not be able to replicate in eukaryotic cells. This minimizes the risk of integration of polynucleotide vaccine sequences into the recipients' genome. Tissue-specific promoters or enhancers may be used whenever it is desirable to limit expression of the polynucleotide to a particular tissue type.

In one embodiment of this invention, the pre-plasmids (e.g., pMRKAd5pol, pMRKAd5nef and pMRKAd5gag were generated by homologous recombination using the MRKHVE3 (and MRKHVO for the E3- version) backbones and the appropriate shuttle vector, as shown for pMRKAd5pol in Figure 22 and for pMRKAd5nef in Figure 23. The plasmid in linear form is capable of replication after entering the PER.C6® cells and virus is produced. The infected cells and media were harvested after viral replication was complete.

10

15

20

25

30

35

Viral vectors can be propagated in various E1 complementing cell lines, including the known cell lines 293 and PER.C6<sup>®</sup>. Both these cell lines express the adenoviral E1 gene product. PER.C6<sup>®</sup> is described in WO 97/00326 (published January 3, 1997) and issued U.S. Patent No. 6,033,908, both of which are hereby incorporated by reference. It is a primary human retinoblast cell line transduced with an E1 gene segment that complements the production of replication deficient (FG) adenovirus, but is designed to prevent generation of replication competent adenovirus by homologous recombination. Cells of particular interest have been stably transformed with a transgene that encodes the AD5E1A and E1B gene, like PER.C6<sup>®</sup>, from 459 bp to 3510 bp inclusive. 293 cells are described in Graham et al., 1977 J. Gen. Virol 36:59-72, which is hereby incorporated by reference. As stated above, consideration must be given to the adenoviral sequences present in the complementing cell line used. It is important that the sequences not overlap with that present in the vector if the possibility of recombination is to be minimized.

It has been found that vectors generated in accordance with the above description are more effective in inducing an immune response and, thus, constitute very promising vaccine candidates. More particularly, it has been found that first generation adenoviral vectors in accordance with the above description carrying a codon-optimized HIV gag gene, regulated with a strong heterologous promoter can be used as human anti-HIV vaccines, and are capable of inducing immune responses.

Standard techniques of molecular biology for preparing and purifying DNA constructs enable the preparation of the DNA immunogens of this invention.

A vaccine composition comprising an adenoviral vector in accordance with the instant invention may contain physiologically acceptable components, such as buffer, normal saline or phosphate buffered saline, sucrose, other salts and polysorbate. One preferred formulation has: 2.5-10 mM TRIS buffer, preferably about 5 mM TRIS buffer; 25-100 mM NaCl, preferably about 75 mM NaCl; 2.5-10% sucrose, preferably about 5% sucrose; 0.01 -2 mM MgCl<sub>2</sub>; and 0.001%-0.01% polysorbate 80 (plant derived). The pH should range from about 7.0-9.0, preferably about 8.0. One skilled in the art will appreciate that other conventional vaccine excipients may also be used it make the formulation. The preferred formulation contains 5mM TRIS, 75 mM NaCl, 5% sucrose, 1mM MgCl<sub>2</sub>, 0.005% polysorbate 80 at pH 8.0 This has a pH and divalent cation composition which is near the optimum for Ad5 stability and minimizes the potential for adsorption of virus to a glass surface. It does not cause tissue irritation upon intramuscular injection. It is preferably frozen until use.

The amount of adenoviral particles in the vaccine composition to be introduced into a vaccine recipient will depend on the strength of the transcriptional and translational promoters used and on the immunogenicity of the expressed gene product. In general, an immunologically or prophylactically effective dose of  $1\times10^7$  to  $1\times10^{12}$  particles and preferably about  $1\times10^{10}$  to  $1\times10^{11}$  particles is administered directly into muscle tissue. Subcutaneous injection, intradermal introduction, impression through the skin, and other modes of administration such as intraperitoneal, intravenous, or inhalation delivery are also contemplated. It is also contemplated that booster vaccinations are to be provided. Following vaccination with HIV adenoviral vector, boosting with a subsequent HIV adenoviral vector and/or plasmid may be desirable. Parenteral administration, such as intravenous, intramuscular, subcutaneous or other means of administration of interleukin-12 protein, concurrently with or subsequent to parenteral introduction of the vaccine compositions of this invention is also advantageous.

The adenoviral vector and/or vaccine plasmids of this invention polynucleotide may be unassociated with any proteins, adjuvants or other agents which impact on the recipients' immune system. In this case, it is desirable for the vector to be in a physiologically acceptable solution, such as, but not limited to, sterile saline or sterile buffered saline. Alternatively, the vector may be associated with an adjuvant known in the art to boost immune responses (i.e., a "biologically effective"

adjuvant), such as a protein or other carrier. Vaccine plasmids of this invention may, for instance, be delivered in saline (e.g., PBS) with or without an adjuvant. Preferred adjuvants are Alum or CRL1005 Block Copolymer. Agents which assist in the cellular uptake of DNA, such as, but not limited to, calcium ions, may also be used to advantage. These agents are generally referred to herein as transfection facilitating reagents and pharmaceutically acceptable carriers. Techniques for coating microprojectiles coated with polynucleotide are known in the art and are also useful in connection with this invention.

10

15

20

25

30

35

This invention also includes a prime and boost regimen wherein a first adenoviral vector is administered, then a booster dose is given. The booster dose may be repeated at selected time intervals. Alternatively, a preferred inoculation scheme comprises priming with a first adenovirus serotype and then boosting with a second adenovirus serotype. More preferably, the inoculation scheme comprises priming with a first adenovirus serotype and then boosting with a second adenovirus serotype, wherein the first and second adenovirus serotypes are classified within separate subgroups of adenoviruses. The above prime/boost schemes are particularly preferred in those situations where a preexisting immunity is identified to the adenoviral vector of choice. In this type of scheme, the individual or population of individuals is primed with an adenovirus of a serotype other than that to which the preexisting immunity is identified. This enables the first adenovirus to effectuate sufficient expression of the transgene while evading existing immunity to the second adenovirus (the boosting adenovirus) and, further, allows for the subsequent delivery of the transgene via the boosting adenovirus to be more effective. Adenovirus serotype 5 is one example of a virus to which such a scheme might be desirable. In accordance with this invention, therefore, one might decide to prime with a non-group C adenovirus (e.g., Ad12, a group A adenovirus, Ad24, a group D adenovirus, or Ad35, a group B adenovirus) to evade anti-Ad5 immunity and then boost with Ad5, a group C adenovirus. Another preferred embodiment involves administration of a different adenovirus (including non-human adenovirus) vaccine followed by administration of the adenoviral vaccines disclosed. In the alternative, a viral antigen of interest can be first delivered via a viral vaccine other than an adenovirus-based vaccine, and then followed with the adenoviral vaccine disclosed. Alternative viral vaccines include but are not limited to pox virus and venezuelan equine encephilitis virus.

A large body of human and animal data supports the importance of cellular immune responses, especially CTL in controlling (or eliminating) HIV infection. In humans, very high levels of CTL develop following primary infection and correlate

with the control of viremia. Several small groups of individuals have been described who are repeatedly exposed to HIV by remain uninfected; CTL has been noted in several of these cohorts. In the SIV model of HIV infection, CTL similarly develops following primary infection, and it has been demonstrated that addition of anti-CD8 monoclonal antibody abrogated this control of infection and leads to disease progression. This invention uses adenoviral vaccines alone or in combination with plasmid vaccines to induce CTL.

The following non-limiting Examples are presented to better illustrate the invention.

10

15

20

25

30

35

5

#### **EXAMPLE 1**

Removal of the Intron A Portion of the hCMV Promoter GMP grade pVIInsHIVgag was used as the starting material to amplify the hCMV promoter. PVIInsHIVgag is a plasmid comprising the CMV immediate-early (IE) promoter and intron A, a full-length codon-optimized HIV gag gene, a bovine growth hormone-derived polyadenylation and transcriptional termination sequence, and a minimal pUC backbone; see Montgomery et al., supra for a description of the plasmid backbone. The amplification was performed with primers suitably positioned to flank the hCMV promoter. A 5' primer was placed upstream of the Msc1 site of the hCMV promoter and a 3' primer (designed to contain the BgIII recognition sequence) was placed 3' of the hCMV promoter. The resulting PCR product (using high fidelity Tag polymerase) which encompassed the entire hCMV promoter (minus intron A) was cloned into TOPO PCR blunt vector and then removed by double digestion with Msc1 and Bg/II. This fragment was then cloned back into the original GMP grade pV1InsHIVgag plasmid from which the original promoter, intron A, and the gag gene were removed following Msc1 and Bg/II digestion. This ligation reaction resulted in the construction of a hCMV promoter (minus intron A) + bGHpA expression cassette within the original pV1JnsHIVgag vector backbone. This vector is designated pVIInsCMV(no intron).

The FLgag gene was excised from pV1JnsHIVgag using BgIII digestion and the 1,526 bp gene was gel purified and cloned into pV1JnsCMV(no intron) at the BgIII site. Colonies were screened using Sma1 restriction enzymes to identify clones that carried the Flgag gene in the correct orientation. This plasmid, designated pV1JnsCMV(no intron)-FLgag-bGHpA, was fully sequenced to confirm sequence integrity.

Two additional transgenes were also constructed. The plasmid, pV1JnsCMV(no intron)-FLgag-SPA, is identical to pV1JnsCMV(no intron)-FLgag-bGHpA except that the bovine growth hormone polyadenylation signal has been replaced with a short synthetic polyA signal (SPA) of 50 nucleotides in length. The sequence of the SPA is as shown, with the essential components (poly(A) site, (GT)<sub>n</sub>, and (T)<sub>n</sub>; respectively) underlined:

<u>AATAAA</u>AGATCTTTATTTTCATTAGATCTGTGTGTTTTTTTTTGTGTG

The plasmid, pV1Jns-mCMV-FLgag-bGHpA, is identical to the pV1JnsCMV(no intron)-FLgag-bGHpA except that the hCMV promoter has been removed and replaced with the murine CMV (mCMV) promoter.

Figure 3 diagrammatically shows the new transgene constructs in comparison with the original transgene.

15

20

25

10

5

(SEQ ID NO:18).

#### **EXAMPLE 2**

Gag Expression Assay for Modified Gag Transgenes

Gag Elisa was performed on culture supernatants obtained from transient tissue culture transfection experiments in which the two new hCMV-containing plasmid constructs, pV1JnsCMV(no intron)-FLgag-bGHpA and pV1JnsCMV(no intron)-FLgag-SPA, both devoid of intron A, were compared to pV1JnsHIVgag which, as noted above possesses the intron A as part of the hCMV promoter. Table 2 below shows the *in vitro* gag expression data of the new gag plasmids compared with the GMP grade original plasmid. The results displayed in Table 2 show that both of the new hCMV gag plasmid constructs have expression capacities comparable to the original plasmid construct which contains the intron A portion of the hCMV promoter.

Table 2: In vitro DNA transfection of original and new plasmid HIV-1 gag constructs.

| Plasmid                              | µg gag/10e6 COS cells/5µg DNA/48 hr |
|--------------------------------------|-------------------------------------|
| HIVFL-gagPR9901 <sup>a</sup>         | 10.8                                |
| PVIIns-hCMV-FLgag-bGHpAb             | 16.6                                |
| pV1Jns-hCMV-FLgag-SPA <sup>b,c</sup> | 12.0                                |

<sup>&</sup>lt;sup>a</sup> GMP grade pV1Jns-hCMVintronA-FLgag-bGHpA.

10

# EXAMPLE 3

Rodent (Balb/c) Study for Modified gag Transgenes
A rodent study was performed on the two new plasmid constructs
described above – pV1JnsCMV(no intron)-FLgag-bGHpA and pV1JnsCMV(no
intron)-FLgag-SPA - in order to compare them with the construct described above
possessing the intron A portion of the CMV promoter, pV1JnsHIVgag. Gag antibody
and Elispot responses (described in PCT International Application No.
PCT/US00/18332 (WO 01/02607) filed July 3, 2000, claiming priority to U.S.
Provisional Application Serial No. 60/142,631, filed July 6, 1999 and U.S.
Application Serial No. 60/148,981, filed August 13, 1999, all three applications which
are hereby incorporated by reference) were measured. The results displayed in Table
3 below, show that the new plasmid constructs behaved equivalently to the original
construct in Balb/c mice with respect to their antibody and T-cell responses at both
dosages of plasmid DNA tested, 20 µg and 200 µg.

b New plasmid constructions that have the intron A portion removed from the hCMV promoter.

<sup>&</sup>lt;sup>c</sup> In this construct the bGH terminator has been replaced with the short synthetic polyadenylation signal (SPA)

**EXAMPLE 4** 

Table 3: HIV191: Immunogenicity of V1Jns-gag under different promoter and termination control elements.

| DNA®                | Dose,<br>ug <sup>b</sup> |       | Anti-p24 Titers<br>(3 Wk PD1) <sup>c</sup> |      |       | SFC/10^6 Cells<br>(4 Wk PD1) <sup>d</sup> |        |
|---------------------|--------------------------|-------|--|------|-------|---|--------|
| Promoter/terminator |                          | GMT   | +SE  | -SE  | Media | gag197-205                                | p24    |
| HIVFL-gagPR9901     | 200                      | 12800 | 4652                                       | 3412 | 2(2)  | 129(19)                                   | 30(11) |
| (GMP grade)         | 20                       | 5572  | 1574                                       | 1227 | Ò     | 56(9)                                     | 25(6)  |
| pV1Jns-hCMV-        | 200                      | 11143 | 2831                                       | 2257 | 0     | 98(5)                                     | 12(6)  |
| FL-gag-bGHpA        | 20                       | 7352  | 2808                                       | 2032 | 0     | 73(9)                                     | 11(6)  |
| pV1Jns-hCMV-        | 200                      | 16890 | 5815                                       | 4326 | 1(1)  | 94(4)                                     | 26(7)  |
| FL-gag-SPA          | 20                       | 5971, | 5361                                       | 2825 | o o   | 85(17)                                    | 38(10) |
| Naïve               | 0                        | 123   | 50   | 36   | 0     | 0   | 0      |

in PBS

20

Construction of the Modified Shuttle Vector - "MRKpdelE1 Shuttle"

The modifications to the original Ad5 shuttle vector (pdelE1sp1A; a vector comprising Ad5 sequences from basepairs 1-341 and 3524-5798, with a multiple cloning region between nucleotides 341 and 3524 of Ad5, included the following three manipulations carried out in sequential cloning steps as follows:

- (1) The left ITR region was extended to include the *Pac1* site at the junction between the vector backbone and the adenovirus left ITR sequences. This allow for easier manipulations using the bacterial homologous recombination system.
- 10 (2) The packaging region was extended to include sequences of the wild-type (WT) adenovirus from 342 bp to 450 bp inclusive.
  - (3) The area downstream of pIX was extended 13 nucleotides (i.e., nucleotides 3511-3523 inclusive).

These modifications (Figure 4) effectively reduced the size of the E1 deletion without overlapping with any part of the E1A/E1B gene present in the transformed PER.C6® cell line. All manipulations were performed by modifying the Ad shuttle vector pdelE1sp1A.

Once the modifications were made to the shuttle vector, the changes were incorporated into the original Ad5 adenovector backbones (pAdHVO and pAdHVE3) by bacterial homologous recombination using *E. coli* BJ5183 chemically competent cells.

bi.m. Injections into both quads, 50 µL per quad

cn=10;GMT, geometric mean titer; SE, standard. error

dn=5, pooled spicens; mean of triplicate wells and standard, deviation, in parentheses;

#### EXAMPLE 5

## Construction of Modified Adenovector Backbones (E3+ and E3-)

5

10

15

20

25

30

The original adenovectors pAdHVO (comprising all Ad5 sequences except those nucleotides encompassing the E1 and E3 regions) and pADHVE3 (comprising all Ad5 sequences except those nucleotides encompassing the E1 region), were each reconstructed so that they contained the modifications to the E1 region. This was accomplished by digesting the newly modified shuttle vector (MRKpdelE1 shuttle) with Pac1 and BstZ1101 and isolating the 2,734 bp fragment which corresponds to the adenovirus sequence. This fragment was co-transformed with DNA from either Cla1 linearized pAdHVO (E3- adenovector) or Cla1 linearized pAdHVE3 (E3+adenovector) into E. coli BJ5183 competent cells. At least two colonies from each transformation were selected and grown in Terrific™ broth for 6-8 hours until turbidity was reached. DNA was extracted from each cell pellet and then transformed into E. coli XL1 competent cells. One colony from each transformation was selected and grown for plasmid DNA purification. The plasmid was analyzed by restriction digestions to identify correct clones. The modified adenovectors were designated MRKpAdHVO (E3- plasmid) and MRKpAdHVE3 (E3+ plasmid). Virus from these new adenovectors (MRKHVO and MRKHVE3, respectively) as well as the old version of the adenovectors were generated in the PER.C6® cell lines to accommodate the following series of viral competition experiments. In addition, the multiple cloning site of the original shuttle vector contained ClaI, BarnHI, Xho I, EcoRV, HindIII, Sal I, and Bgl II sites. This MCS was replaced with a new MCS containing Not I, Cla I, EcoRV and Asc I sites. This new MCS has been transferred to the MRKpAdHVO and MRKpAdHVE3 pre-plasmids along with the modification made to the packaging region and pIX gene.

#### EXAMPLE 6

# Analysis of the Effect of the Packaging Signal Extension

To study the effects of the modifications made to the E1 deletion region, the viruses obtained from the original backbone (pAdHVE3) and the new backbone (MRKpAdHVE3) were mixed together in equal MOI ratios (1:1 and 5:5) and passaged through several rounds; see Figure 5, Expt.#1. Both of the viruses in the experiment contained the E3 gene intact and did not contain a transgene. The only difference between the two viruses was within the region of the E1 deletion.

Following the coinfection of the viruses at P1 (passage 1), the mixtures were propagated through an additional 4 passages at which time the cells were harvested

and the virus extracted and purified by CsCl banding. The viral DNA was extracted and digested with *Hind*III and the digestion products were then radioactively labeled. For the controls, the respective pre-plasmids (pAdHVE3 ("OLD E3+"); MRKpAdHVE3 ("NEW E3+")) were also digested with *Hind*III (and *Pac1* to remove the vector backbone) and subsequently labeled with [<sup>33</sup>P]dATP. The radioactively labeled digestion products were subjected to gel electrophoresis and the gel was dried down onto Whatman paper before being exposed to autoradiographic film. Figure 6 clearly shows that the new adenovirus which has the addition made to the packaging signal region has a growth advantage compared with the original adenovirus. In the experiments performed (at either ratio tested), only the digestion bands pertaining to the newly modified virus were present. The diagnostic band of size 3,206 (from the new virus) was clearly present. However, there was no evidence of the diagnostic band of size 2,737 bp expected from the original virus.

15

20

25

30

35

10

#### EXAMPLE 7

# Analysis of the Effect of the E3 Gene

The second set of the virus competition study involved mixing equal MOI ratio (1:1) of the newly modified viruses, that obtained from MRKpAdHVO and MRKpAdHVE3 (Figure 5, Expt. #2). In this set, both viruses had the new modifications made to the E1 deletion. The first virus (that from MRKpAdHVO) does not contain an E3 gene. The second virus (that from MRKpAdHVE3) does contain the E3 gene. Neither of the viruses contain a transgene. Following coinfection of the viruses, the mixtures were propagated through an additional 4 passages at which time the cells were harvested and the total virus extracted and purified by CsCl banding. The viral DNA was extracted and digested with HindIII and the digestion products were then radioactively labeled. For the controls, the respective pre-plasmids MRKpAdHVO ("NEW E3-"); MRKpAdHVE3 ("NEW E3+") were also digested with HindIII (and Pac1 to remove the vector backbone) and then labeled with [33P]dATP. The radioactively labeled digestion products were subjected to gel electrophoresis and the gel was dried down onto Whatman paper before being exposed to autoradiographic film. Figure 6 shows the results of the viral DNA analysis of the E3+ virus and E3- virus mixing experiment. The diagnostic band corresponding to the E3+ virus (5,665 bp) was present in greater amount compared with the diagnostic band of 3,010 bp corresponding to the E3- virus. This indicates that the virus that contains the E3 gene is able to amplify more rapidly

compared with the virus that does not contain an E3 gene. This increased amplification capacity has been confirmed by growth studies; see Table 4 below.

#### **EXAMPLE 8**

# Construction of the new shuttle vector containing modified gag transgene – "MRKpdelE1-CMV(no intron)-FLgag-bGHpA"

The modified plasmid pV1JnsCMV(no intron)-FLgag-bGHpA was digested with Msc1 overnight and then digested with Sfi1 for 2 hours at 50°C. The DNA was then treated with Mungbean nuclease for 30 mins at 30°C. The DNA mixture was desalted using the Qiaex II kit and then Klenow treated for 30 mins at 37°C to fully blunt the ends of the transgene fragment. The 2,559 bp transgene fragment was then gel purified. The modified shuttle vector (MRKpdelE1 shuttle) was linearized by digestion with EcoRV, treated with calf intestinal phosphatase and the resulting 6,479 bp fragment was then gel purified. The two purified fragments were then ligated together and several dozen clones were screened to check for insertion of the transgene within the shuttle vector. Diagnostic restriction digestion was performed to identify those clones carrying the transgene in the E1 parallel and E1 anti-parallel orientation. This strategy was followed to clone in the other gag transgenes in the MRKpdelE1 shuttle vector.

20

5

10

15

#### EXAMPLE 9

# Construction of the MRK FG Adenovectors

The shuttle vector containing the HIV-1 gag transgene in the E1 parallel orientation, MRKpdelE1-CMV(no intron)-FLgag-bGHpA, was digested with Pac1. 25 The reaction mixture was digested with BsfZ171. The 5,291 bp fragment was purified by gel extraction. The MRKpAdHVE3 plasmid was digested with Cla1 overnight at 37°C and gel purified. About 100 ng of the 5,290 bp shuttle +transgene fragment and ~100 ng of linearized MRKpAdHVE3 DNA were co-transformed into E. coli BJ5183 chemically competent cells. Several clones were selected and grown in 2 ml Terrific™ broth for 6-8 hours, until turbidity was reached. The total DNA from the 30 cell pellet was purified using Qiagen alkaline lysis and phenol chloroform method. The DNA was precipitated with isopropanol and resuspended in 20 µl dH<sub>2</sub>0. A 2 µl aliquot of this DNA was transformed into E. coli XL-1 competent cells. A single colony from each separate transformation was selected and grown overnight in 3 ml 35 LB +100 μg/ml ampicillin. The DNA was isolated using Qiagen columns. A positive clone was identified by digestion with the restriction enzyme BstEII which cleaves

within the gag gene as well as the plasmid backbone. The pre-plasmid clone is designated MRKpAdHVE3+CMV(no intron)-FLgag-bGHpA and is 37,498 bp in size. This strategy was followed to generate E3- and E3+ versions of each of the other gag transgene constructions in both E1 parallel and E1 anti-parallel versions. Figures 7A, 7B and 7C show the various combinations of adenovectors constructed.

# **EXAMPLE 10**

# Plasmid Competition Studies

A series of plasmid competition studies was carried out. Briefly, the screening of the various combinations of new constructs was performed by mixing equal amounts of each of two competing plasmids. In the experiment shown in Figure 8A, plasmids containing the same transgene but in different orientations were mixed together to create a "competition" between the two plasmids. The aim was to look at the effects of transgene orientation. In the experiment shown in Figure 8B, plasmids containing different polyadenylation signals (but in the same orientation) were mixed together in equal amounts. The aim was to assess effects of polyA signals. Following the initial transfection, the virus was passaged through ten rounds and the viral DNA analyzed by radioactive restriction analysis.

Analysis of the viral species from the plasmid mixing experiment (Figure 8A) showed that adenovectors which had the transgene inserted in the E1 parallel orientation amplified better and were able to out-compete the adenovirus which had the transgene inserted in the E1 anti-parallel orientation. Viral DNA analysis of the mixtures at passage 3 and certainly at passage 6, showed a greater ratio of the virus carrying the transgene in the E1 parallel orientation compared with the E1 antiparallel version. By passage 10, the only viral species observed was the adenovector with the transgene in the E1 parallel orientation for both transgenes tested (hCMV(no intron)-FLgag-bGHpA and hCMV(no intron)-FLgag-SPA).

Analysis of the viral species from the plasmid mixing experiment #2 (Figure 8B) at passages 3 and 6 showed that the polyadenylation signals tested (bGHpA and SPA) did not have an effect on the growth of the virus. Even at passage 10 the two viral species in the mixture were still present in equal amounts.

. 30

5

10

15

20

25

#### **EXAMPLE 11**

Virus generation of an enhanced adenoviral construct - "MRK Ad5 HIV-1gag"

The results obtained from the competition study allowed us to make the following conclusions: (1) The packaging signal extension is beneficial; (2) Presence of E3 does enhance viral growth; (3) E1 parallel orientation is recommended; and (4) PolyA signals have no effect on the growth of the adenovirus.

MRK Ad5 HIV-1 gag exhibited the most desirable results. This construct contains the hCMV(no intron)-FLgag-bGHpA transgene inserted into the new E3+adenovector backbone, MRKpAdHVE3, in the E1 parallel orientation. We have designated this adenovector MRK Ad5 HIV-1 gag. This construct was prepared as outlined below:

The pre-plasmid MRKpAdHVE3+CMV(no intron)-FLgag-bGHpA was digested was Pac1 to release the vector backbone and 3.3 µg was transfected by calcium phosphate method (Amersham Pharmacia Biotech.) in a 6 cm dish containing PER.C6<sup>®</sup> cells at ~60% confluence. Once CPE was reached (7-10 days), the culture was freeze/thawed three times and the cell debris pelleted. 1 ml of this cell lysate was used to infect into a 6 cm dish containing PER.C6<sup>®</sup> cells at 80-90% confluence. Once CPE was reached, the culture was freeze/thawed three times and the cell debris pelleted. The cell lysate was then used to infect a 15 cm dish containing PER.C6<sup>®</sup> cells at 80-90% confluence. This infection procedure was continued and expanded at passage 6. The virus was then extracted from the cell pellet by CsCl method. Two bandings were performed (3-gradient CsCl followed by a continuous CsCl gradient). Following the second banding, the virus was dialyzed in A105 buffer. Viral DNA was extracted using pronase treatment followed by phenol chloroform. The viral DNA was then digested with *Hind*III and radioactively labeled with [33P]dATP. Following gel electrophoresis to separate the digestion products the gel was dried down on Whatman paper and then subjected to autoradiography. The digestion products were compared with the digestion products from the pre-plasmid (that had been digested with Pac1/HindIII prior to labeling). The expected sizes were observed, indicating that the virus had been successfully rescued. This strategy was used to rescue virus from each of the various adenovector plasmid constructs prepared.

5

10

15

20

25

30

#### EXAMPLE 12

# Stability Analyses

5

10

15

25

30

35

To determine whether the various adenovector constructs (e.g., MRK Ad5 HIV-1 gag) show genetic stability, the viruses were each passaged continually. The viral DNA was analyzed at passages 3, 6 and 10. Each virus maintained its correct genetic structure. In addition, the stability of the MRK Ad5 HIV-1 gag was analyzed under propagation conditions similar to that performed in large scale production. For this analysis, the transfections of MRK Ad5 HIV-1 gag as well as three other adenoviral vectors were repeated and the virus was purified at P3. The three other adenovectors were as follows: (1) that comprising hCMV(no intron)-Flgag with a bGHpA terminator in an E3- adenovector backbone; (2) that comprising hCMV(no intron)-Flgag with a SPA termination signal in an E3+ adenovector backbone, and that comprising a mCMV-Flgag with a bGHpA terminator in an E3+ adenovector backbone. All of the vectors have the transgene inserted in the E1 parallel orientation. Viral DNA was analyzed by radioactive restriction analysis to confirm that it was correct before being delivered to fermentation cell culture for continued passaging in serum-free media. At P5 each of the four viruses were purified and the viral DNA extracted for analysis by the restriction digestion and radiolabeling procedure. This virus has subsequently been used in a series of studies (in vitro gag expression in COS cells, rodent study and rhesus monkey study) as will be described below. The viruses 20 from P5 are shown in Figure 9.

The passaging under serum-free conditions was continued for the MRKHVE3 (transgene-less, obtained from MRKpAdHVE3 pre-plasmid) and the MRKAd5HIV-1gag (obtained from MRKpAdHVE3+CMV(no intron)-FLgagbGHpA pre-plasmid) viruses. Figure 10 shows viral DNA analysis by radioactive restriction digestion at passage 11 for MRKHVE3, MRKAd5HIV-1gagE3-, and passage 11 and 12 for MRKAd5HIV-1gag. Aside from the first lane which is the DNA marker lane, the next three lanes are virus from the pre-plasmid controls (controls based on the original virus) - MRKpAdHVE3 (also referred to as "pMRKHVE3"), MRKpAdHVE3+CMV(no intron)-FLgag-bGHpA, and pMRKAd5gag(E3-), respectively. As seen in Figure 10, each of the viral DNA samples show the expected bands with no extraneous bands showing. This signifies that there are no major variant adenovirus species present that can be detected by autoradiography.

Figure 11 shows the results of viral competition study between MRKHVE3 and MRKAd5HIV-1 gag. These viruses were mixed together at equal MOI (140 viral

particles each; 280 vp total) at passage 6 and continued to be passaged until P11. Aside from the first lane which is the DNA marker lane, the next two lanes are the pre-plasmid controls obtained from MRKpAdHVE3 and MRKpAdHVE3+CMV(no intron)-FLgag-bGHpA. The next two lanes are the viral DNA from the starting viral material at passage six. The last two lanes are the competition studies performed in duplicate. The data in Figure 11 shows the effect the gag transgene in culture. Growth of a MRKAd5gag virus was compared with growth of a "transgene-less" MRKHVE3. These two viruses were infected at the same MOI (i.e. 140 vp each) at passage 6 and then passaged through to passage 11 and the viral pool was analyzed by radioactive restriction analysis. The data shows that one virus did not out compete the other. Therefore, the gag transgene did not show obvious signs of toxicity to the adenovirus.

10

15

20

25

30

Analysis by *Hind*III digestion shows that each virus specie is present in approximately equal amounts. As above, there does not appear to be signs of any extraneous bands. Figure 12 shows higher passage numbers for MRKAd5HIV-1gag grown under serum-containing conditions. The genome integrity again has been maintained and there is no evidence of rearrangements, even at the highest passage level (P21).

Each of the four vectors shown in Figure 9 were analyzed for amplification capacity. Table 4 below shows the QPA analysis used in the estimation of viral amplification ratios at P4. The determination of the amplification ratio for the original HIV-1 gag construct is based on the clinical lot at P12. It has been shown that amplification rates increases with higher passage number for the original virus. The reason for this observation is due to the emergence of variants which exhibit increased growth rates compared to the intact adenovector. With continued passaging of the original Ad gag vector, the level of variants increases and hence amplification rates increase also.

The MRK Ad5 HIV-1 gag virus has also been continually passaged under process conditions (i.e., serum-free media). Viral DNA extracted from passages 11 and 12 show no evidence of rearrangement.

Table 4:
Amplification Ratios Based on AEX and QPA Analysis of Virus Amplification from Passage 3 to Passage 4.

|                        | ·                   |  |  |
|------------------------|---------------------|--|--|
| Ad gag construct       | Amplification Ratio |  |  |
| MRKAd5gag              | 470                 |  |  |
| HCMV-Flgag-bGHpA [E3-] | 115                 |  |  |
| HCMV-Flgag-SPA [E3+]   | 320                 |  |  |
| mCMV-FLgag-bGHpA [E3+] | 420                 |  |  |
| Original construct *   | 40 - 50             |  |  |

5

10

15

20

25

#### EXAMPLE 13

Analytical Evaluation of the enhanced Ad5 Constructs

To study the effects of the transgene and the E3 gene on virus amplification, the enhanced adenoviral vector, MRK Ad5 HIV-1 gag, along with its transgene-less version (MRKpAdHVE3) and its E3- version (MRK Ad5 HIV-1 gag E3-), was studied for several passages under serum-free conditions. Table 5A shows the amplification ratios determined for passages P3 to P8 for MRK Ad5 HIV-1 gag. Within a certain MOI range, it has been determined that the virus output is directly proportional to the virus input. Therefore, the greater the number of virus particles per cell at infection, the greater the virus amount produced. Viral amplification ratios, on the other hand, are inversely proportional to the virus input. The lower the virus input, the greater the amplification ratio.

Table 5B shows the amplification rates of the new E3+ vector backbone MRKpAdHVE3. It has a significantly lower rate of amplification compared with the gag transgene containing version. This may be contributed to the larger size MRK Ad5 HIV-1 gag since it contains the transgene. This inclusion of the transgene brings the size of the adenovirus closer to the size of a wild type Ad5 virus. It is well known that adenoviruses amplify best when they are at close to their wild type genomic size.

<sup>\*</sup> This estimation is based on the clinical lot growth characteristics at Passage 12.

Wild type Ad5 is 35,935 bp. The MRKpAdHVE3 is 32, 905 bp in length. The enhanced adenovector MRK Ad5 HIV-1 gag is 35,453bp (See Figure 14 for vector map; see also Figure 15A-X show the complete pre-adenoviral vector sequence, which includes an additional 2,021 bp of the vector backbone).

5

10

Table 5C shows the amplification rates of the new E3- gag containing virus MRK Ad5 HIV-1 gag E3-. Once again, this virus shows lower growth rate than the enhanced adenoviral vector. This may be attributed to the decreased sized of this virus (due to the E3 gene deletion) compared with wild type Ad5. The MRK Ad5 HIV-1 gag E3- virus is 32,810 bp in length. This can be compared with the wild type Ad5 which is 35,935 bp and MRK Ad5 HIV-1 gag which is 35,453 bp in length.

**Table 5A:** Amplification ratios determined by AEX and QPA for MRKAd5gag over several continuous passaging in serum free media. Following P5, two replicate samples were taken (rep-1 and rep-2) and analyzed.

### MRKAd5gag rep1

|     | XV (10° calls/m<br>Intection | D, Viablity (%)<br>Harvest | Harvest Time | Cell Pessage<br>Number | Tiber<br>10 <sup>00</sup> vp/mi culture | Titer<br>10° vp/cet | OPA<br>10° TCID <sub>eo</sub> fent | Ratio<br>AEXCOPA | Ampideadon<br>Ratio | AEX<br>Internal Control |
|-----|------------------------------|----------------------------|--------------|------------------------|---|---------------------|------------------------------------|------------------|---------------------|-------------------------|
| P4  | 1,49,81%                     | 0.58, 50%                  | 44           | 46                     | 8.7                                     | 5.9                 | 1.72                               | 50               | 470<br>(MOI = 125)  |                         |
| P5  | 1,38,93%                     | 0.65, 47%                  | 48           | 49                     | 6.7                                     | 4.9                 | 1.38                               | 49               | 170                 |                         |
| PB  | 1,04, 94%                    | 0.68, 77%                  | 47           | 48                     | 5.8                                     | 5.6                 | 1.42                               | 41               | 200                 |                         |
| P7  | 1.50, 84%                    | 0.95, 61%                  | 49.5         | 50                     | 3.9                                     | 1.4                 | 0.97                               | 40               | 50                  | 1                       |
| P7  | 1.09, 97%                    | 0.76, 59%                  | 50           | 52                     | 5.2                                     | 4.7                 | 1.70                               | 81               | 170                 | 1                       |
| P8  | 1.03, 94%                    | 0.86, 64%                  | 47.5         | 64                     | 8.0                                     | 8.7                 | 1.10                               | 62               | 310                 | 1                       |
| P9  | 0,69, 95%                    | 0.99, 73%                  | 47.5         | 58                     | 4,4                                     | 4.9                 | 1.03                               | 43               | 175                 | 3.12<br>2.84            |
| P10 | 1.09, 91%                    | 1.06, 65%                  | 47.5         | 58                     | 8.0                                     | 2.8                 | 1.16                               | 26               | 100                 | 2.70<br>2.60            |
| P11 | 1.19, 88%                    | 0.98, 65%                  | 47           | 60                     | 3.6                                     | 3.0                 | 1.15                               | 31               | 110                 | 2.70<br>2.70            |
| P12 | 0.58, 91%                    | 0.85, 63%                  | 47.5         | 47                     | 5.4                                     | 6.5                 | 1.20                               | 45               | 200                 | 2.88<br>2.60            |
| P13 | 1,00, 88%                    | 0.70, 67%                  | 49           | 49                     | 5.8                                     | 5.8                 | 1.11                               | 52               | 210                 | 3.18<br>3.18            |
| P14 | 1.94, 92%                    | 0.88, 67%                  | 46 .         | 53                     | 8.6                                     | 4.4                 |                                    |                  | 160                 | 3.2B<br>3.27            |
| P15 | 0.57, 96%                    | 0.64, 66%                  | 47           | 47                     | 6.9                                     | 7.1                 |                                    |                  | 250                 | 3.12<br>2.91            |

**Table 5B:** Amplification ratios determined by AEX and QPA for MRKHVE3 over several continuous passaging in serum free media. MRKHVE3 is the new vector backbone which does NOT carry a transgene.

#### MRKHVE3

|     | Xv (10° cells/n | ni), Viability (%)<br>Harvest | Harvest Time | Cell Passage<br>Number | Titer<br>10 <sup>10</sup> vp/ml culture | Titler<br>10° vp/cett | QPA<br>10" TCID <sub>80</sub> /ml | Ratio<br>AEX:QPA | Amplification<br>Ratio | AEX<br>Internal Control |
|-----|-----------------|-------------------------------|--------------|------------------------|---|-----------------------|-----------------------------------|------------------|------------------------|-------------------------|
| P4  | 1.10, 97%       | 1.28, 79%                     | 49           | 54                     | 4.1                                     | 3.8                   | 1.70                              | 25               | 300<br>(MO) = 125)     |                         |
| P5  | 0.92, 69%       | 1.18, 77%                     | 47           | . 48                   | 4.3                                     | 4.7                   | 1.24                              | 35               | 170                    |                         |
| P6  | 1.55, 86%       | 1,28, 76%                     | 49.5         | 50                     | 12                                      | 0.8                   | 0.58                              | 21               | 30                     |                         |
| P6  | 1.09, 97%       | 1.11,81%                      | 49           | 52                     | 4.0                                     | 3.6                   | 1.16                              | 34               | 130                    |                         |
| P7  | 1.17, 91%       | 1,22, 91%                     | 47.5         | 54                     | 3.7                                     | 3.2                   | 0.50                              | 74               | 110                    |                         |
| P8  | 0.98, 88%       | 1,41, 83%                     | 48           | 58                     | 21                                      | 21                    | 0.47                              | 45               | 75                     | 3.12<br>2.64            |
| P9  | 1,20, 89%       | 1.26, 81%                     | 47.5         | 58                     | 0.8                                     | 0.7                   | 0.29                              | 28               | 25                     | 2.70<br>2.60            |
| PID | 0.99, 82%       | 1.65, 85%                     | 47           | 50                     | 23                                      | 23                    | 0.43                              | . 53             | 80                     | 2.70<br>2.70            |
| PII | 1,07,96%        | 1.25, 83%                     | 48           | 47                     | 2.7                                     | 2.5                   | 0.41                              | 66               | 90                     | 2.86<br>2.60            |
| P12 | 0.80, 91%       | 1.14, 80%                     | 49.5         | 49                     | 5.9                                     | 7.4                   | 0.48                              | 123              | 260                    | 3.18<br>3.18            |
| P13 | 1.95, 95%       | 1.14, 85%                     | 45.5         | 53                     | 5.8                                     | 3.0                   | <u> </u>                          |                  | 110                    | 3.28<br>3.27            |
| P14 | 0.97, 95%       | 1.03, 98%                     | 48.5         | 47                     | 9,4                                     | 8.7                   |                                   |                  | 350                    | 3.12<br>2.91            |
| P15 | 0.87, 99%       | 0.97, 89%                     | 49.5         | 49                     | 5.3                                     | 6.1                   |                                   |                  | 218                    | 2.78<br>2.52            |

Table 5C. Amplification ratios determined by AEX and QPA for MRKAd5gag(E3-) over several continuous passaging in serum free media. This construct is identical to the MRKAd5gag construct except that this version is DELETED of the E3 gene.

# 5

10

15

#### MRKAd5gag(E3-)

|     |           | ni), Viebliny (%) | Harvest Time | Cell Passage | Titer                          | Titer       | QPA                        | Ratio   | Amplification    | AEX                  |
|-----|-----------|-------------------|--------------|--------------|--------------------------------|-------------|----------------------------|---------|------------------|----------------------|
|     | Infection | Harvest           | h.p.l.       | Number       | 10 <sup>th</sup> vp/ml culture | 10° vp/ceti | 10° TCID <sub>so</sub> Ami | AEX:QPA | Ratio            | Internal Contro      |
| P4  | 1.62, 77% | 1.12, 62%         | 47.5         | 46           | 2.0                            | 1.2         | 0,92                       | 20      | 100<br>(MOI=125) | ·                    |
| P5  | 1.16, 92% | 0.62, 43%         | 49           | 49           | 3.3                            | 2.9         | 0.99                       | 34      | 100              |                      |
| P6  | 1.71, 86% | 0.20, 10%         | 49           | 50           | 4.7                            | 2.7         | 1.70                       | 28      | 100              |                      |
| P6  | 1.09, 97% | 0.63, 54%         | 49.5         | 52           | 5.4                            | 5.0         | 1.76                       | 31      | 180              |                      |
| P7  | 1.17, 91% | 0.98, 72%         | 47.50        | 54           | 7.1                            | 6.1         | 0.67                       | 106     | 220              |                      |
| P8  | 0.98, 68% | 0.77, 48%         | 48           | 56           | 8.1                            | 3.2         | 0.66                       | 47      | 115              | 3.12                 |
| P9  | 1.20, 89% | 1.03, 72%         | 48           | 58           | 1.8                            | 1.5         | 0.57                       | 32      | 55               | 2.84<br>2.70<br>2.60 |
| P10 | 0.99, 82% | 0.80, 62%         | 46.5         | 60           | 3.2                            | 3.2         | 0.68                       | 47      | 115              | 2.70                 |
| P11 | 1.07, 96% | 0.88, 70%         | 48.5         | 47           | 5.9                            | 5.5         | 0.68                       | 87      | 200              | 2.70<br>2.88<br>2.60 |
| P12 | 0.80, 91% | 0.57, 59%         | 50           | 49           | 5.1                            | 6.4         | 0.72                       | 71      | 230              | 3.18<br>3.18         |
| P13 | 1.96, 95% | 0.91, 59%         | 45.5         | 53           | 7.4                            | 3.8         | <u> </u>                   |         | 135              | 3.28                 |
| P14 | 0.97, 96% | 0.81, 74%         | 48           | 47           | 6.8                            | 7.0         | l                          |         | 250              | 3.27<br>3.12         |
| P16 | 0.87, 99% | 0.84, 56%         | 49           | 49           | 4.8                            | 5.5         |                            |         | 196              | 2.91<br>2.78<br>2.52 |

#### **EXAMPLE 14**

### Gag Expression Analysis of the Novel Constructs

In vitro gag analysis of the MRK Ad5 HIV-1 gag and the original HIV-gag vectors (research and clinical lot) show comparable gag expression. The clinical lot shows only a slightly reduced gag expression level. The most noticeable difference is with the mCMV vector. This vector shows roughly 3 fold lower expression levels compared with the other vectors tested (which all contain hCMV promoters). The mCMV-FLgag with bGHpA assay was performed three times using different propagation and purification lots and it consistently exhibited weaker gag expression.

#### EXAMPLE 15

# Evaluation of MRK Ad5 HIV-1 gag and Other gag-Containing Adenovectors in Balb/c Mice

Cohorts of 10 balb/c mice were vaccinated intramuscularly with escalating doses of MRK Ad5 HIV-1 gag, and the research and clinical lots of original Ad5HIV-1gag. Serum samples were collected 3 weeks post dose 1 and analyzed by anti-p24 sandwich ELISA.

Anti-p24 titers in mice that received MRK Ad5 HIV-1 gag (107 and 109 vp(viral particle) doses) were comparable (Figure 13) to those of the research lot of Ad5HIV-1 gag, for which much of the early rhesus data were generated on. These titers were also comparable when E3 is deleted (MRKAd5hCMVgagbGHpA(E3-)) or SPA is substituted for bGHpA terminator (MRKAd5 hCMV-gag-SPA (E3+)) or murine CMV promoter is used in place of hCMV (MRKAd5 mCMV-gag-bGHpA (E3+)) in the MRKAd5 backbone.

The results shown in Table 7 indicate that the three other vectors (in addition to the preferred vector, MRK Ad5 HIV-1 gag, are also capable of inducing strong anti-gag antibody responses in mice. Interestingly enough, while the mCMV-FLgag construct containing bGHpA and E3+ in an E1 parallel orientation showed lowest gag expression in the COS cell *in vitro* infection (Table 6) in comparison with the other vectors tested, it generated the greatest anti-gag antibody response this *in vivo* Balb/c study. Table 7 also shows a dose response in anti-gag antibody production in both the research and the clinical lot. As expected, the clinical lot shows reduced anti-gag antibody induction at each dosage level compared to the same dosage used for the research lot.

Table 6: In vitro analysis for gag expression in COS cells by Elisa assay.

15

10

| Viral Vectors <sup>a</sup>       | µg gag/4.8x10e5 COS/10e8 parts/48hr |
|----------------------------------|-------------------------------------|
| MRKAd5gag <sup>b</sup>           | 1.40                                |
| Clinical lot Ad5gag <sup>c</sup> | 1.28                                |
| Research lot Ad5gag <sup>d</sup> | 1.32                                |
| MCMVFL-gagbGHpA <sup>c</sup>     | 0.42                                |

<sup>&</sup>lt;sup>a</sup> A<sub>260mm</sub> absorbance readings taken for viral particle determinations.

<sup>&</sup>lt;sup>b</sup> MRKAd5gag was produced in serum free conditions and purified at P5.

<sup>&</sup>lt;sup>c</sup> Clinical lot# Ad5gagFN0001

<sup>25</sup> dResearch Ad5FLgag lot# 6399

<sup>°</sup> mCMVFL-gagbGHpA was produced in serum free conditions and purified at P5.

Table 7: mHIV020 Anti-p24 Ab Titers in Balb/c mice (n=10) vaccinated with various Adgag constructs and lots (3 week post dose1).

| Group<br>ID | Vaccine   | Dose<br>(vp) | GMT    | SE upper | SE lower |
|-------------|---|--------------|--------|----------|----------|
| 1           | <sup>a</sup> MRKAd5gag                          | 10^7         | 25600  | 5877     | 4780     |
| 2           | n n   | 10^9         | 409600 | 94028    | 76473    |
| 3           | hCMV FL-gag bGHpA [E3-] →                       | 10^7         | 7352   | 2077     | 1620     |
| 4           | ,   | 10^9         | 235253 | 59767    | 47659    |
| 5           | hCMV FL-gag SPA [Ē3+] →                         | 10^7         | 12800  | 9905     | 236      |
| 6           | •   | 10^9         | 310419 | 99181    | 75165    |
| 7           | bmCMV FL-gag bGHpA [E3+] →                      | 10^7         | 44572  | 23504    | 15389    |
| 8           | •   | 10^9         | 941014 | 239068   | 190836   |
| 9           | <sup>c</sup> hCMV FL-gag bGHpA [E3-] ←          | 10^7         | 3676   | 934      | 745      |
| 10          |   | 10^9         | 117627 | 17491    | 15227    |
| 11          | research lot hCMV intronA FL-gag bGHpA [E3-] <- | 10^6         | 528    | 262      | 175      |
| 12          |   | 10^7         | 14703  | 5274     | 3882     |
| 13          | <u> </u>  | 10^8         | 58813  | 14942    | 11915    |
| 14          |   | 10^9         | 204800 | 53232    | 42250    |
| 15          | clinical lot hCMVintronA FL-gag bGHpA [E3-] <-  | 10^6         | 230    | 82       | 61       |
| 16          |   | 10^7         | 4222   | 3405     | 1138     |
| 17          |   | 10^8         | 19401  | 3939     | 3274     |
| 18          | ) H   | 10^9         | 89144  | 25187    | 19639    |
| 19          | Naĭve   | none         | 93     | 7        | 6        |

\*2x50 µL i.m. (quad) injections/animal .

P.I.s: Youil, Chen, Casimiro Vaccination: T. Toner, Q. Su

Assay: M. Chen

5

10

<sup>a</sup>The structure of MRKAd5gag is: hCMVFL-gagbGHpA [E3+] → The <u>same lot</u> of MRKAd5gag used in this rodent study was used in the Rhesus monkey study (Tables 7 and 8).

#### **EXAMPLE 16**

Comparison of Humoral and Cellular Responses Towards the Original Ad-gag Construct with the New MRK Ad5 HIV-1 gag in Rhesus Monkeys

- Cohorts of 3 rhesus monkeys were vaccinated intramuscularly with MRK Ad5 HIV-1 gag or the clinical Ad5gag bulk at two doses,  $10^{11}$  vp and  $10^9$  vp. Immunizations were conducted at week 0, 4, and 25. Serum and PBMC samples were collected at selected time points. The serum sample were assayed for anti-p24 Ab titers (using competitive based assay) and the PBMCs for antigen-specific IFN-gamma secretion following overnight stimulation with gag 20-mer peptide pool (via ELISpot assay).
- The results shown in Table 8 indicate comparable responses with respect to the generation of anti-gag antibodies. The frequencies of gag-specific T cells in

The same lot of mCMVFL-gagbGHpA[E3+] used in the *in vitro* study (Table 6) ws used here.

<sup>&</sup>lt;sup>c</sup>This construct was designed by Volker Sandig. It contains a shorter version of the hCMV promoter than that used in the MRK constructs. The adenovector backbone is identical to the original backbone used in the original Adgag vector. Expression at 10e7 dose from this vector is 7 fold lower then the same dose of the MRKAd5gag and 4 fold lower than the research lot.

peripheral blood assummarized in Table 9 demonstrate a strong cellular immune response generated after a single dose with the new construct MRK Ad5 HIV-1 gag. The responses are also boostable with second dose of the same vector. The vector is also able to induce CD8+ T cell responses (as evident by remaining spot counts after CD4+ depletion of PBMCs) which are responsible for cytotoxic activity.

Table 8 Anti-p24 antibody titers (in mMU/mL) in rhesus macaques immunized with

gag-expressing adenovectors (Protocol HIV203).

| Vaccine                                      | Pre      | Wk4           | Wk 8_    | Wk 12   | Wk 16 | Wk 20    | Wk 25    | Wk 28    |
|--|----------|---------------|----------|---------|-------|----------|----------|----------|
| MRKAd5gag°, 10^11 vp                         |          |               |          |         |       |          | ·        |          |
| 97N010                                       | <10      | 118           | 5528     | 11523   | 7062  | 21997    | ND       | 51593    |
| 97N116                                       | <10      | 62            | 772      | 1447    | 1562  | 2174     | ND       | 20029    |
| 98X007                                       | <10      | 66            | 3353     | 6156    | 6845  | 3719     | ND_      | 24031    |
| MRKAd5gog, 10^9 vp                           |          |               |          |         |       |          |          |          |
| 97N120                                       | <10      | 51            | 204      | 318     | 366   | 482      | ND       | 6550     |
| 97N144                                       | <10      | 18            | 118      | 274     | 706   | 888      | ND       | 7136     |
| 98X008                                       | <10      | 15            | 444      | 386     | 996   | 1072     | ND       | 12851    |
| Ad5gag <sup>b</sup> , Clinical Lat, 10^11 vp |          |               |          |         |       |          |          |          |
| 97X001                                       | <10      | 87            | 2579_    | 4718    | 7174  | 7250     | ND       | 69226    |
| 97N146                                       | <10      | 72            | 3604_    | 7380    | 7526  | 18906    | ND       | 60283    |
| 98X009                                       | <10      | 78            | 4183     | 3946    | 3124  | 6956     | ND       | 26226    |
| Ad5gag, Clinical Lot, 10^9 vp                |          |               |          |         |       |          |          |          |
| 97N020                                       | <10      | <10           | 143      | 371     | 390   | 1821     | ND       | 17177    |
| 97X003                                       | <10      | <10           | 39       | 93      | 156   | 596      | ND       | 2053     |
| 98X012                                       | <10      | 81            | 342      | 717     | 956   | 1558     | ND       | 11861    |
| MRKAdagag (hCMV, bGHpA, E3+)                 |          |               |          | <b></b> |       | <u> </u> | <u> </u> | <u> </u> |
| Pariginal Actigag vector (hCMV/Intro         | n A bGHp | A. E3-), lott | #FN0001_ |         | L     | L        | <u> </u> |          |
| ND, not determined                           |          |               |          |         |       |          | <u> </u> |          |

Table 9. Number of gag-specific T cells per million peripheral blood mononuclear cells (PBMCs) in rhesus monkeys immunized with gag-expressing adenovectors. Also included are those frequencies in PBMCs depleted of CD4<sup>+</sup> T cells.

| Grp # | Vaccination         | Monkey ID    |       | Wk                 |       | Wk    |       | Wk    | I=10  | Wk     | T=2   | . Wk | T=2   | Wk    |
|-------|---------------------|--------------|-------|--------------------|-------|-------|-------|-------|-------|--------|-------|------|-------|-------|
|       | T=0,4,25 w/s        |              | Medic | Gog H <sup>b</sup> | Media | Gog H | Medio | Gog H | Media | Goog H | Media | GogH | Media | Gog H |
|       |                     |              |       |                    |       |       |       |       |       | _      |       |      |       |       |
| 1     | MRKAc5ccc           | 97ND1D       | В     | 89                 | 0     | 395   | 0     | 1058  | 0     | 1174   | 3     | 775  | 4     | 1074  |
|       | 1041 VP             | 97N010(CD4-) | 4     | 38                 | 1     |       | 3     | 993   |       |        | 0     | 76   | 0     | 594   |
|       |                     | 97N116       | 1     | 398                | 1     | 609   | 0     | 534   | 4     | 395    | 1     | 261  | 0     | 408   |
|       |                     | 97N116(CD4-) | 11    | 676                | · ·   |       | 0     | 593   |       |        | 0     | 184  | 0     | 666   |
|       |                     | 98X007       | 10    | 579                | 0     | 1304  | 3     | 2193  | 1     | 2118   | 3     | 1588 | 0     | 2113  |
|       |                     | 98X007(CD4-) | 20    | 965                |       |       | 0     | 2675  |       |        | 0     | 1656 | 0     | 1278  |
| 2     | MRKAdacco           | 97N120       | 5     | 275                | 1     | 249   | 4     | 141   | 4     | 119    | 9     | 206  | 4     | 219   |
|       | 10/9 VP             | 97N120(CD4-) | 111   | 170                |       | l .   | 0     | 85    |       | 1      | 0     | 75   | 1     | 219   |
|       |                     | 97N144       | 3     | 235                | 6     | 438   | 1     | 318   | 3     | 256    | 1     | 98   | 5     | 373   |
|       |                     | 97N144(CD4-) | 6     | 148                | 1     |       | 0     | 285   | l .   |        | ND    | NO   | 10    | 625   |
|       |                     | 98X008       | 4     | 368                | 1     | 1090  | 3     | 891   | 4     | 673    | 3     | 473  | 5     | 735   |
|       |                     | 98X008(CD4-) | 14    | 696                |       | 1     | 0     | 1175  | 1     | 1      | ٥     | 391  | 4     | 848   |
| 3     | AdSpag clinical lat | 97X001       | В     | 261                | 1     | 485   | 0     | 817   | 0     | 1220b  | 1     | 894  | 0     | 1858  |
| ٠.    | 10^11 vp            | 97X001(CD4-) | 10    | 283                | j ·   |       | 3     | 996   |       | 1      | 0     | 1010 | 0     | 1123  |
|       | •                   | 97N146       | 3     | 150                | 1 1   | 465   | 0     | 339   | 1     | 1272   | 3     | 1238 | 3     | 1785  |
|       |                     | 97N146(CD4-) | 6     | 133                | 1     | Į.    | 0     | 370   | 1     | l      | 0     | 654  | 0     | 971   |
|       |                     | 9800009      | 0     | 93                 | 3     | 339   | ] 3   | 559   | 0     | 896    | 1     | 384  | 0     | 1748  |
|       |                     | 98X009(CD4-) | Ö     | 73                 |       |       | 0     | 333   |       | 1      | ٥     | 225  | 0     | 644   |
| 4     | AdSgoog dinical lat | 97N020       | 3     | 30                 | 1     | 101   | 0     | 66    | 0     | 36     | 0     | 26   | 0     | 41    |
|       | 10/9 vp             | 97N020(CD4-) | 10    | 29                 | 1 .   |       | 0     | 15    | l .   | l      | 0     | 1 1  | [ 0   | 16    |
|       |                     | 97X003       | 4     | 68                 | 5     | 134   | 0     | 18    | 1 1   | 38     | 4     | 38   | 6     | 81    |
|       |                     | 97X003(CD4-) | ٩     | 40                 | 1     |       | 0     | 6     | 1 .   | 1      | 0     | 4    | 0     | 19    |
|       |                     | 98X012       | 5     | 95                 | 3     | 54    | 1 1   | 34    | 0     | 18     | 0     | 20   | 1 1   | 121   |
|       |                     | 98X012(CD4-) | 111   | 70                 |       | }     | ٥     | 11    |       | 1      | ٥     | B    | 0     | 41    |
| 5     | Nave                | 96R041       | 6     | 8                  | 1     | 1     | 000   | 0     | 0     | 0      | 0     | 0    | 1     | 0     |
|       |                     | 053F         | 14    | 18                 | 5     | 16.   | 20    | 14    | 19    | 15     | 10    | 15   | 24    | °     |

Based on either 4x10/5 or 2x10/5 cells per well (depending on spot density)

ND, not determined "mock or no peolide control

10

15

20

<sup>b</sup>Pool of 20-capeatides overlapping by 10 aband encorposising the passequence

The adenovectors described herein and, particularly, MRK Ad5 HIV-1 gag, represent very promising HIV-gag adenovectors with respect to their enhanced growth characteristics in both serum and, more importantly, in serum-free media conditions. In comparison with the current HIV-1 gag adenovector construct, MRK Ad5 HIV-1 gag shows a 5-10 fold increased amplification rate. We have shown that it is genetically stable at passage 21. This construct is able to generate significant cellular immune responses in vivo even at a relatively low dose of 10^9 vp. The potency of the MRKAd5gag construct is comparable to, if not better than the original HIV-1gag vector as shown in this rhesus monkey study.

# EXAMPLE 17 CODON OPTIMIZED HIV-1 POL AND CODON OPTIMIZED HIV-1 POL MODIFICATIONS

The open reading frames for the various synthetic *pol* genes disclosed herein comprise coding sequences for the reverse transcriptase (or RT which consists of a polymerase and RNase H activity) and integrase (IN). The protein sequence is based

5

10

15

20

25

30

35

on that of Hxb2r, a clonal isolate of IIIB; this sequence has been shown to be closest to the consensus clade B sequence with only 16 nonidentical residues out of 848 (Korber, et al., 1998, Human retroviruses and AIDS, Los Alamos National Laboratory, Los Alamos, New Mexico). The skilled artisan will understand after review of this specification that any available HIV-1 or HIV-2 strain provides a potential template for the generation of HIV pol DNA vaccine constructs disclosed herein. It is further noted that the protease gene is excluded from the DNA vaccine constructs of the present invention to insure safety from any residual protease activity in spite of mutational inactivation. The design of the gene sequences for both wildtype (wt-pol) and inactivated pol (IA-pol) incorporates the use of human preferred ("humanized") codons for each amino acid residue in the sequence in order to maximize in vivo mammalian expression (Lathe, 1985, J. Mol. Biol. 183:1-12). As can be discerned by inspecting the codon usage in SEQ ID NOs: 1, 3, 5 and 7, the following codon usage for mammalian optimization is preferred: Met (ATG), Gly (GGC), Lys (AAG), Trp (TGG), Ser (TCC), Arg (AGG), Val (GTG), Pro (CCC), Thr (ACC), Glu (GAG); Leu (CTG), His (CAC), Ile (ATC), Asn (AAC), Cys (TGC), Ala (GCC), Gln (CAG), Phe (TTC) and Tyr (TAC). For an additional discussion relating to mammalian (human) codon optimization, see WO 97/31115 (PCT/US97/02294), which, as noted elsewhere in this specification, is hereby incorporated by reference. It is intended that the skilled artisan may use alternative versions of codon optimization or may omit this step when generating HIV pol vaccine constructs within the scope of the present invention. Therefore, the present invention also relates to non-codon optimized versions of DNA molecules and associated recombinant adenoviral HIV vaccines which encode the various wild type and modified forms of the HIV Pol protein disclosed herein. However, codon optimization of these constructs is a preferred embodiment of this invention.

A particular embodiment of this portion of the invention comprisies codon optimized nucleotide sequences which encode wt-pol DNA constructs (herein, "wt-pol" or "wt-pol (codon optimized))" wherein DNA sequences encoding the protease (PR) activity are deleted, leaving codon optimized "wild type" sequences which encode RT (reverse transcriptase and RNase H activity) and IN integrase activity. A DNA molecule which encodes this protein is disclosed herein as SEQ ID NO:1, the open reading frame being contained from an initiating Met residue at nucleotides 10-12 to a termination codon from nucleotides 2560-2562. SEQ ID NO:1 is as follows:

AGATCTACCA TGGCCCCCAT CTCCCCCATT GAGACTGGC CTGTGAAGCT GAAGCCTGGC

ATGGATGGCC CCAAGGTGAA GCAGTGGCCC CTGACTGAGG AGAAGATCAA GGCCCTGGTG

|    | GAAATCTGCA | CTGAGATGGA | GAAGGAGGC  | AAAATCTCCA | AGATTGGCCC | CGAGAACCCC |
|----|------------|------------|------------|------------|------------|------------|
|    | TACAACACCC | CTGTGTTTGC | CATCAAGAAG | AAGGACTCCA | CCAAGTGGAG | GAAGCTGGTG |
|    | GACTTCAGGG | AGCTGAACAA | GAGGACCCAG | GACTTCTGGG | AGGTGCAGCT | GGGCATCCCC |
|    | CACCCCCCTG | GCCTGAAGAA | GAAGAAGTCT | GTGACTGTGC | TGGATGTGGG | GGATGCCTAC |
| 5  | TTCTCTGTGC | CCCTGGATGA | GGACTTCAGG | AAGTACACTG | CCTTCACCAT | CCCCTCCATC |
|    | AACAATGAGA | CCCCTGGCAT | CAGGTACCAG | TACAATGTGC | TGCCCCAGGG | CTGGAAGGGC |
|    | TCCCCTGCCA | TCTTCCAGTC | CTCCATGACC | AAGATCCTGG | AGCCCTTCAG | GAAGCAGAAC |
|    | CCTGACATTG | TGATCTACCA | GTACATGGAT | GACCTGTATG | TGGGCTCTGA | CCTGGAGATT |
|    | GGGCAGCACA | GGACCAAGAT | TGAGGAGCTG | AGGCAGCACC | TGCTGAGGTG | GGGCCTGACC |
| 10 | ACCCCTGACA | AGAAGCACCA | GAAGGAGCCC | CCCTTCCTGT | GGATGGGCTA | TGAGCTGCAC |
|    | CCCGACAAGT | GGACTGTGCA | GCCCATTGTG | CTGCCTGAGA | AGGACTCCTG | GACTGTGAAT |
|    | GACATCCAGA | AGCTGGTGGG | CAAGCTGAAC | TGGGCCTCCC | AAATCTACCC | TGGCATCAAG |
|    | GTGAGGCAGC | TGTGCAAGCT | GCTGAGGGGC | ACCAAGGCCC | TGACTGAGGT | GATCCCCCTG |
|    | ACTGAGGAGG | CTGAGCTGGA | GCTGGCTGAG | AACAGGGAGA | TCCTGAAGGA | GCCTGTGCAT |
| 15 | GGGGTGTACT | ATGACCCCTC | CAAGGACCTG | ATTGCTGAGA | TCCAGAAGCA | GGGCCAGGGC |
|    | CAGTGGACCT | ACCAAATCTA | CCAGGAGCCC | TTCAAGAACC | TGAAGACTGG | CAAGTATGCC |
|    | AGGATGAGGG | GGGCCCACAC | CAATGATGTG | AAGCAGCTGA | CTGAGGCTGT | GCAGAAGATC |
|    | ACCACTGAGT | CCATTGTGAT | CTGGGGCAAG | ACCCCCAAGT | TCAAGCTGCC | CATCCAGAAG |
|    | GAGACCTGGG | AGACCTGGTG | GACTGAGTAC | TGGCAGGCCA | CCTGGATCCC | TGAGTGGGAG |
| 20 | TTTGTGAACA | CCCCCCCCT  | GGTGAAGCTG | TGGTACCAGC | TGGAGAAGGA | GCCCATTGTG |
|    | GGGGCTGAGA | CCTTCTATGT | GGATGGGGCT | GCCAACAGGG | AGACCAAGCT | GGGCAAGGCT |
|    | GGCTATGTGA | CCAACAGGGG | CAGGCAGAAG | GTGGTGACCC | TGACTGACAC | CACCAACCAG |
|    | AAGACTGAGC | TCCAGGCCAT | CTACCTGGCC | CTCCAGGACT | CTGGCCTGGA | GGTGAACATT |
|    | GTGACTGACT | CCCAGTATGC | CCTGGGCATC | ATCCAGGCCC | AGCCTGATCA | GTCTGAGTCT |
| 25 | GAGCTGGTGA | ACCAGATCAT | TGAGCAGCTG | ATCAAGAAGG | AGAAGGTGTA | CCTGGCCTGG |
|    | GTGCCTGCCC | ACAAGGGCAT | TGGGGGCAAT | GAGCAGGTGG | ACAAGCTGGT | GTCTGCTGGC |
|    | • •        | -          |            |            | AGGATGAGCA |            |
|    | CACTCCAACT | GGAGGGCTAT | GGCCTCTGAC | TTCAACCTGC | CCCCTGTGGT | GGCTAAGGAG |
|    | ATTGTGGCCT | CCTGTGACAA | GTGCCAGCTG | AAGGGGGAGG | CCATGCATGG | GCAGGTGGAC |
| 30 | TGCTCCCCTG | GCATCTGGCA | GCTGGACTGC | ACCCACCTGG | AGGGCAAGGT | GATCCTGGTG |
|    | GCTGTGCATG | TGGCCTCCGG | CTACATTGAG | GCTGAGGTGA | TCCCTGCTGA | GACAGGCCAG |
|    | GAGACTGCCT | ACTTCCTGCT | GAAGCTGGCT | GGCAGGTGGC | CTGTGAAGAC | CATCCACACT |
|    |            |            |            |            |            | GGCTGGCATC |
|    | AAGCAGGAGT | TTGGCATCCC | CTACAACCCC | CAGTCCCAGG | GGGTGGTGGA | GTCCATGAAC |
| 35 | AAGGAGCTGA | AGAAGATCAT | TGGGCAGGTG | AGGGACCAGG | CTGAGCACCT | GAAGACAGCT |
|    | GTGCAGATGG | CTGTGTTCAT | CCACAACTTC | AAGAGGAAGG | GGGGCATCGG | GGGCTACTCC |

GCTGGGGAGA GGATTGTGGA CATCATTGCC ACAGACATCC AGACCAAGGA GCTCCAGAAG
CAGATCACCA AGATCCAGAA CTTCAGGGTG TACTACAGGG ACTCCAGGAA CCCCCTGTGG
AAGGGCCCTG CCAAGCTGCT GTGGAAGGGG GAGGGGGCTG TGGTGATCCA GGACAACTCT
GACATCAAGG TGGTGCCCAG GAGGAAGGCC AAGATCATCA GGGACTATGG CAAGCAGATG
GCTGGGGATG ACTGTGTGGC CTCCAGGCAG GATGAGGACT AAAGCCCGGG CAGATCT (SEQ
ID NO:1).

The open reading frame of the wild type pol construct disclosed as SEQ ID NO:1 contains 850 amino acids, disclosed herein as SEQ ID NO:2, as follows: Met Ala Pro Ile Ser Pro Ile Glu Thr Val Pro Val Lys Leu Lys Pro Gly Met Asp Gly Pro Lys Val Lys Gln Trp Pro Leu Thr Glu Glu Lys 10 Ile Lys Ala Leu Val Glu Ile Cys Thr Glu Met Glu Lys Glu Gly Lys Ile Ser Lys Ile Gly Pro Glu Asn Pro Tyr Asn Thr Pro Val Phe Ala Ile Lys Lys Lys Asp Ser Thr Lys Trp Arg Lys Leu Val Asp Phe Arg Glu Leu Asn Lys Arg Thr Gln Asp Phe Trp Glu Val Gln Leu Gly Ile Pro His Pro Ala Gly Leu Lys Lys Lys Lys Ser Val Thr Val Leu Asp 15 Val Gly Asp Ala Tyr Phe Ser Val Pro Leu Asp Glu Asp Phe Arg Lys Tyr Thr Ala Phe Thr Ile Pro Ser Ile Asn Asn Glu Thr Pro Gly Ile Arg Tyr Gln Tyr Asn Val Leu Pro Gln Gly Trp Lys Gly Ser Pro Ala Ile Phe Gln Ser Ser Met Thr Lys Ile Leu Glu Pro Phe Arg Lys Gln Asn Pro Asp Ile Val Ile Tyr Gln Tyr Met Asp Asp Leu Tyr Val Gly 20 Ser Asp Leu Glu Ile Gly Gln His Arg Thr Lys Ile Glu Glu Leu Arg Gln His Leu Leu Arg Trp Gly Leu Thr Thr Pro Asp Lys Lys His Gln Lys Glu Pro Pro Phe Leu Trp Met Gly Tyr Glu Leu His Pro Asp Lys Trp Thr Val Gln Pro Ile Val Leu Pro Glu Lys Asp Ser Trp Thr Val 25 Asn Asp Ile Gln Lys Leu Val Gly Lys Leu Asn Trp Ala Ser Gln Ile Tyr Pro Gly Ile Lys Val Arg Gln Leu Cys Lys Leu Leu Arg Gly Thr Lys Ala Leu Thr Glu Val Ile Pro Leu Thr Glu Glu Ala Glu Leu Glu Leu Ala Glu Asn Arg Glu Ile Leu Lys Glu Pro Val His Gly Val Tyr Tyr Asp Pro Ser Lys Asp Leu Ile Ala Glu Ile Gln Lys Gln Gly Gln Gly Gln Trp Thr Tyr Gln Ile Tyr Gln Glu Pro Phe Lys Asn Leu Lys 30 Thr Gly Lys Tyr Ala Arg Met Arg Gly Ala His Thr Asn Asp Val Lys Gln Leu Thr Glu Ala Val Gln Lys Ile Thr Thr Glu Ser Ile Val Ile Trp Gly Lys Thr Pro Lys Phe Lys Leu Pro Ile Gln Lys Glu Thr Trp Glu Thr Trp Trp Thr Glu Tyr Trp Gln Ala Thr Trp Ile Pro Glu Trp Glu Phe Val Asn Thr Pro Pro Leu Val Lys Leu Trp Tyr Gln Leu Glu 35 Lys Glu Pro Ile Val Gly Ala Glu Thr Phe Tyr Val Asp Gly Ala Ala

Asn Arg Glu Thr Lys Leu Gly Lys Ala Gly Tyr Val Thr Asn Arg Gly Arg Gln Lys Val Val Thr Leu Thr Asp Thr Thr Asn Gln Lys Thr Glu Leu Gln Ala Ile Tyr Leu Ala Leu Gln Asp Ser Gly Leu Glu Val Asn Ile Val Thr Asp Ser Gln Tyr Ala Leu Gly Ile Ile Gln Ala Gln Pro Asp Gln Ser Glu Ser Glu Leu Val Asn Gln Ile Ile Glu Gln Leu Ile Lys Lys Glu Lys Val Tyr Leu Ala Trp Val Pro Ala His Lys Gly Ile Gly Gly Asn Glu Gln Val Asp Lys Leu Val Ser Ala Gly Ile Arg Lys Val Leu Phe Leu Asp Gly Ile Asp Lys Ala Gln Asp Glu His Glu Lys Tyr His Ser Asn Trp Arg Ala Met Ala Ser Asp Phe Asn Leu Pro Pro 10 Val Val Ala Lys Glu Ile Val Ala Ser Cys Asp Lys Cys Gln Leu Lys Gly Glu Ala Met His Gly Gln Val Asp Cys Ser Pro Gly Ile Trp Gln Leu Asp Cys Thr His Leu Glu Gly Lys Val Ile Leu Val Ala Val His Val Ala Ser Gly Tyr Ile Glu Ala Glu Val Ile Pro Ala Glu Thr Gly Gln Glu Thr Ala Tyr Phe Leu Leu Lys Leu Ala Gly Arg Trp Pro Val Lys Thr Ile His Thr Asp Asn Gly Ser Asn Phe Thr Gly Ala Thr Val 15 Arg Ala Ala Cys Trp Trp Ala Gly Ile Lys Gln Glu Phe Gly Ile Pro Tyr Asn Pro Gln Ser Gln Gly Val Val Glu Ser Met Asn Lys Glu Leu Lys Lys Ile Ile Gly Gln Val Arg Asp Gln Ala Glu His Leu Lys Thr Ala Val Gln Met Ala Val Phe Ile His Asn Phe Lys Arg Lys Gly Gly 20 Ile Gly Gly Tyr Ser Ala Gly Glu Arg Ile Val Asp Ile Ile Ala Thr Asp Ile Gln Thr Lys Glu Leu Gln Lys Gln Ile Thr Lys Ile Gln Asn Phe Arg Val Tyr Tyr Arg Asp Ser Arg Asn Pro Leu Trp Lys Gly Pro Ala Lys Leu Leu Trp Lys Gly Glu Gly Ala Val Val Ile Gln Asp Asn Ser Asp Ile Lys Val Val Pro Arg Arg Lys Ala Lys Ile Ile Arg Asp 25 Tyr Gly Lys Gln Met Ala Gly Asp Asp Cys Val Ala Ser Arg Gln Asp Glu Asp (SEQ ID NO:2).

The present invention especially relates to an adenoviral vector vaccine which comprises a codon optimized HIV-1 DNA pol construct wherein, in addition to deletion of the portion of the wild type sequence encoding the protease activity, a combination of active site residue mutations are introduced which are deleterious to HIV-1 pol (RT-RH-IN) activity of the expressed protein. Therefore, the present invention preferably relates to an adenoviral HIV-1 DNA pol-based vaccine wherein the construct is devoid of DNA sequences encoding any PR activity, as well as containing a mutation(s) which at least partially, and preferably substantially, abolishes RT, RNase and/or IN activity. One type of HIV-1 pol mutant which is part and parcel of an adenoviral vector vaccine may include but is not limited to a mutated

30

35

5

10

15

20

25

DNA molecule comprising at least one nucleotide substitution which results in a point mutation which effectively alters an active site within the RT, RNase and/or IN regions of the expressed protein, resulting in at least substantially decreased enzymatic activity for the RT, RNase H and/or IN functions of HIV-1 Pol. In a preferred embodiment of this portion of the invention, a HIV-1 DNA pol construct contains a mutation or mutations within the Pol coding region which effectively abolishes RT, RNase H and IN activity. An especially preferable HIV-1 DNA pol construct in a DNA molecule which contains at least one point mutation which alters the active site of the RT, RNase H and IN domains of Pol, such that each activity is at least substantially abolished. Such a HIV-1 Pol mutant will most likely comprise at least one point mutation in or around each catalytic domain responsible for RT, RNase H and IN activity, respectfully. To this end, an especially preferred HIV-1 DNA pol construct is exemplified herein and contains nine codon substitution mutations which results in an inactivated Pol protein (IA Pol: SEQ ID NO:4, Figure 17A-C) which has no PR, RT, RNase or IN activity, wherein three such point mutations reside within each of the RT, RNase and IN catalytic domains. Therefore, an especially preferred exemplification is an adenoviral vaccine which comprises, in an appropriate fashion, a DNA molecule which encodes IA-pol, which contains all nine mutations as shown below in Table 1. An additional preferred amino acid residue for substitution is Asp551, localized within the RNase domain of Pol. Any combination of the mutations disclosed herein may suitable and therefore may be utilized as an IA-Pol-based vaccine of the present invention. While addition and deletion mutations are contemplated and within the scope of the invention, the preferred mutation is a point mutation resulting in a substitution of the wild type amino acid with an alternative amino acid residue.

| ~ | • | •   | - |
|---|---|-----|---|
|   |   |     |   |
|   |   | 163 |   |

|    |       |            | <del></del> |                 |
|----|-------|------------|-------------|-----------------|
| •  | wt aa | aa residue | mutant aa   | enzyme function |
|    | Asp   | 112        | Ala         | RT              |
|    | Asp   | 187        | Ala         | RT              |
| 30 | Asp   | 188        | Ala         | RT              |
|    | Asp   | 445        | . Ala       | . RNase H       |
|    | Glu   | 480        | Ala         | RNase H         |
|    | Asp   | 500        | Ala         | RNase H         |
|    | Asp   | 626        | Ala         | IN              |
| 35 | Asp   | 678        | Ala         | IN              |
|    | Glu   | 714        | Ala         | IN              |
|    |       |            |             |                 |

It is preferred that point mutations be incorporated into the IApol mutant adenoviral vaccines of the present invention so as to lessen the possibility of altering epitopes in and around the active site(s) of HIV-1 Pol.

To this end, SEQ ID NO:3 discloses the nucleotide sequence which codes for a codon optimized pol in addition to the nine mutations shown in Table 1, disclosed as follows, and referred to herein as "IApol":

10

15

20

25

30

35

AGATCTACCA TGGCCCCCAT CTCCCCCATT GAGACTGTGC CTGTGAAGCT GAAGCCTGGC ATGGATGGCC CCAAGGTGAA GCAGTGGCCC CTGACTGAGG AGAAGATCAA GGCCCTGGTG GAAATCTGCA CTGAGATGGA GAAGGAGGGC AAAATCTCCA AGATTGGCCC CGAGAACCCC TACAACACCC CTGTGTTTGC CATCAAGAAG AAGGACTCCA CCAAGTGGAG GAAGCTGGTG GACTTCAGGG AGCTGAACAA GAGGACCCAG GACTTCTGGG AGGTGCAGCT GGGCATCCCC CACCCCGCTG GCCTGAAGAA GAAGAAGTCT GTGACTGTGC TGGCTGTGGG GGATGCCTAC TTCTCTGTGC CCCTGGATGA GGACTTCAGG AAGTACACTG CCTTCACCAT CCCCTCCATC AACAATGAGA CCCCTGGCAT CAGGTACCAG TACAATGTGC TGCCCCAGGG CTGGAAGGGC TCCCCTGCCA TCTTCCAGTC CTCCATGACC AAGATCCTGG AGCCCTTCAG GAAGCAGAAC CCTGACATTG TGATCTACCA GTACATGGCT GCCCTGTATG TGGGCTCTGA CCTGGAGATT GGGCAGCACA GGACCAAGAT TGAGGAGCTG AGGCAGCACC TGCTGAGGTG GGGCCTGACC ACCCCTGACA AGAAGCACCA GAAGGAGCCC CCCTTCCTGT GGATGGGCTA TGAGCTGCAC CCCGACAAGT GGACTGTGCA GCCCATTGTG CTGCCTGAGA AGGACTCCTG GACTGTGAAT GACATCCAGA AGCTGGTGGG CAAGCTGAAC TGGGCCTCCC AAATCTACCC TGGCATCAAG GTGAGGCAGC TGTGCAAGCT GCTGAGGGGC ACCAAGGCCC TGACTGAGGT GATCCCCCTG ACTGAGGAGG CTGAGCTGGA GCTGGCTGAG AACAGGGAGA TCCTGAAGGA GCCTGTGCAT GGGGTGTACT ATGACCCCTC CAAGGACCTG ATTGCTGAGA TCCAGAAGCA GGGCCAGGGC CAGTGGACCT ACCAAATCTA CCAGGAGCCC TTCAAGAACC TGAAGACTGG CAAGTATGCC AGGATGAGGG GGGCCCACAC CAATGATGTG AAGCAGCTGA CTGAGGCTGT GCAGAAGATC ACCACTGAGT CCATTGTGAT CTGGGGCAAG ACCCCCAAGT TCAAGCTGCC CATCCAGAAG GAGACCTGGG AGACCTGGTG GACTGAGTAC TGGCAGGCCA CCTGGATCCC TGAGTGGGAG TTTGTGAACA CCCCCCCCT GGTGAAGCTG TGGTACCAGC TGGAGAAGGA GCCCATTGTG GGGGCTGAGA CCTTCTATGT GGCTGGGGCT GCCAACAGGG AGACCAAGCT GGGCAAGGCT GGCTATGTGA CCAACAGGGG CAGGCAGAAG GTGGTGACCC TGACTGACAC CACCAACCAG AAGACTGCCC TCCAGGCCAT CTACCTGGCC CTCCAGGACT CTGGCCTGGA GGTGAACATT GTGACTGCCT CCCAGTATGC CCTGGGCATC ATCCAGGCCC AGCCTGATCA GTCTGAGTCT GTGCCTGCCC ACAAGGGCAT TGGGGGCAAT GAGCAGGTGG ACAAGCTGGT GTCTGCTGGC ATCAGGAAGG TGCTGTTCCT GGATGGCATT GACAAGGCCC AGGATGAGCA TGAGAAGTAC CACTCCAACT GGAGGGCTAT GGCCTCTGAC TTCAACCTGC CCCCTGTGGT GGCTAAGGAG

PCT/US01/28861 WO 02/022080

ATTGTGGCCT CCTGTGACAA GTGCCAGCTG AAGGGGGAGG CCATGCATGG GCAGGTGGAC TGCTCCCCTG GCATCTGGCA GCTGGCCTGC ACCCACCTGG AGGGCAAGGT GATCCTGGTG GCTGTGCATG TGGCCTCCGG CTACATTGAG GCTGAGGTGA TCCCTGCTGA GACAGGCCAG GAGACTGCCT ACTTCCTGCT GAAGCTGGCT GGCAGGTGGC CTGTGAAGAC CATCCACACT GCCAATGGCT CCAACTTCAC TGGGGCCACA GTGAGGGCTG CCTGCTGGTG GGCTGGCATC AAGCAGGAGT TTGGCATCCC CTACAACCCC CAGTCCCAGG GGGTGGTGGC CTCCATGAAC AAGGAGCTGA AGAAGATCAT TGGGCAGGTG AGGGACCAGG CTGAGCACCT GAAGACAGCT GTGCAGATGG CTGTGTTCAT CCACAACTTC AAGAGGAAGG GGGGCATCGG GGGCTACTCC GCTGGGGAGA GGATTGTGGA CATCATTGCC ACAGACATCC AGACCAAGGA GCTCCAGAAG CAGATCACCA AGATCCAGAA CTTCAGGGTG TACTACAGGG ACTCCAGGAA CCCCCTGTGG AAGGGCCCTG CCAAGCTGCT GTGGAAGGGG GAGGGGGCTG TGGTGATCCA GGACAACTCT GACATCAAGG TGGTGCCCAG GAGGAAGGCC AAGATCATCA GGGACTATGG CAAGCAGATG GCTGGGGATG ACTGTGTGGC CTCCAGGCAG GATGAGGACT AAAGCCCGGG CAGATCT (SEQ ID NO:3).

. 5

10

15

20

35

In order to produce the IA-pol-based adenoviral vaccines of the present invention, inactivation of the enzymatic functions was achieved by replacing a total of nine active site residues from the enzyme subunits with alanine side-chains. As shown in Table 1, all residues that comprise the catalytic triad of the polymerase, namely Asp112, Asp187, and Asp188, were substituted with alanine (Ala) residues (Larder, et al., Nature 1987, 327: 716-717; Larder, et al., 1989, Proc. Natl. Acad. Sci. 1989, 86: 4803-4807). Three additional mutations were introduced at Asp445, Glu480 and Asp500 to abolish RNase H activity (Asp551 was left unchanged in this IA Pol construct), with each residue being substituted for an Ala residue, respectively (Davies, et al., 1991, Science 252:, 88-95; Schatz, et al., 1989, FEBS Lett. 257: 311-314; Mizrahi, et al., 1990, Nucl. Acids. Res. 18: pp. 5359-5353). HIV pol integrase 25 function was abolished through three mutations at Asp626, Asp678 and Glu714. Again, each of these residues has been substituted with an Ala residue (Wiskerchen, et al., 1995, J. Virol. 69: 376-386; Leavitt, et al., 1993, J. Biol. Chem. 268: 2113-2119). Amino acid residue Pro3 of SEQ ID NO:4 marks the start of the RT gene. The complete amino acid sequence of IA-Pol is disclosed herein as SEQ ID NO:4 and 30 Figure 17A-C, as follows:

Met Ala Pro Ile Ser Pro Ile Glu Thr Val Pro Val Lys Leu Lys Pro Gly Met Asp Gly Pro Lys Val Lys Gln Trp Pro Leu Thr Glu Glu Lys Ile Lys Ala Leu Val Glu Ile Cys Thr Glu Met Glu Lys Glu Gly Lys Ile Ser Lys Ile Gly Pro Glu Asn Pro Tyr Asn Thr Pro Val Phe Ala Ile Lys Lys Lys Asp Ser Thr Lys Trp Arg Lys Leu Val Asp Phe Arg

Glu Leu Asn Lys Arg Thr Gln Asp Phe Trp Glu Val Gln Leu Gly Ile Pro His Pro Ala Gly Leu Lys Lys Lys Ser Val Thr Val Leu Ala Val Gly Asp Ala Tyr Phe Ser Val Pro Leu Asp Glu Asp Phe Arg Lys Tyr Thr Ala Phe Thr Ile Pro Ser Ile Asn Asn Glu Thr Pro Gly Ile Arg Tyr Gln Tyr Asn Val Leu Pro Gln Gly Trp Lys Gly Ser Pro Ala Ile Phe Gln Ser Ser Met Thr Lys Ile Leu Glu Pro Phe Arg Lys Gln Asn Pro Asp Ile Val Ile Tyr Gln Tyr Met Ala Ala Leu Tyr Val Gly Ser Asp Leu Glu Ile Gly Gln His Arg Thr Lys Ile Glu Glu Leu Arg Gln His Leu Leu Arg Trp Gly Leu Thr Thr Pro Asp Lys Lys His Gln Lys Glu Pro Pro Phe Leu Trp Met Gly Tyr Glu Leu His Pro Asp Lys 10 Trp Thr Val Gln Pro Ile Val Leu Pro Glu Lys Asp Ser Trp Thr Val Asn Asp Ile Gln Lys Leu Val Gly Lys Leu Asn Trp Ala Ser Gln Ile Tyr Pro Gly Ile Lys Val Arg Gln Leu Cys Lys Leu Leu Arg Gly Thr Lys Ala Leu Thr Glu Val Ile Pro Leu Thr Glu Glu Ala Glu Leu Glu Leu Ala Glu Asn Arg Glu Ile Leu Lys Glu Pro Val His Gly Val Tyr 15 Tyr Asp Pro Ser Lys Asp Leu Ile Ala Glu Ile Gln Lys Gln Gly Gln Gly Gln Trp Thr Tyr Gln Ile Tyr Gln Glu Pro Phe Lys Asn Leu Lys Thr Gly Lys Tyr Ala Arg Met Arg Gly Ala His Thr Asn Asp Val Lys Gln Leu Thr Glu Ala Val Gln Lys Ile Thr Thr Glu Ser Ile Val Ile 20 Trp Gly Lys Thr Pro Lys Phe Lys Leu Pro Ile Gln Lys Glu Thr Trp Glu Thr Trp Trp Thr Glu Tyr Trp Gln Ala Thr Trp Ile Pro Glu Trp Glu Phe Val Asn Thr Pro Pro Leu Val Lys Leu Trp Tyr Gln Leu Glu Lys Glu Pro Ile Val Gly Ala Glu Thr Phe Tyr Val Ala Gly Ala Ala Asn Arg Glu Thr Lys Leu Gly Lys Ala Gly Tyr Val Thr Asn Arg Gly 25 Arg Gln Lys Val Val Thr Leu Thr Asp Thr Thr Asn Gln Lys Thr Ala Leu Gln Ala Ile Tyr Leu Ala Leu Gln Asp Ser Gly Leu Glu Val Asn Ile Val Thr Ala Ser Gln Tyr Ala Leu Gly Ile Ile Gln Ala Gln Pro Asp Gln Ser Glu Ser Glu Leu Val Asn Gln Ile Ile Glu Gln Leu Ile Lys Lys Glu Lys Val Tyr Leu Ala Trp Val Pro Ala His Lys Gly Ile Gly Gly Asn Glu Gln Val Asp Lys Leu Val Ser Ala Gly Ile Arg Lys 30 . Val Leu Phe Leu Asp Gly Ile Asp Lys Ala Gln Asp Glu His Glu Lys Tyr His Ser Asn Trp Arg Ala Met Ala Ser Asp Phe Asn Leu Pro Pro Val Val Ala Lys Glu Ile Val Ala Ser Cys Asp Lys Cys Gln Leu Lys Gly Glu Ala Met His Gly Gln Val Asp Cys Ser Pro Gly Ile Trp Gln Leu Ala Cys Thr His Leu Glu Gly Lys Val Ile Leu Val Ala Val His Val Ala Ser Gly Tyr Ile Glu Ala Glu Val Ile Pro Ala Glu Thr Gly

Gln Glu Thr Ala Tyr Phe Leu Leu Lys Leu Ala Gly Arg Trp Pro Val Lys Thr Ile His Thr Ala Asn Gly Ser Asn Phe Thr Gly Ala Thr Val Arg Ala Ala Cys Trp Trp Ala Gly Ile Lys Gln Glu Phe Gly Ile Pro Tyr Asn Pro Gln Ser Gln Gly Val Val Ala Ser Met Asn Lys Glu Leu Lys Lys Ile Ile Gly Gln Val Arg Asp Gln Ala Glu His Leu Lys Thr Ala Val Gln Met Ala Val Phe Ile His Asn Phe Lys Arg Lys Gly Gly Ile Gly Gly Tyr Ser Ala Gly Glu Arg Ile Val Asp Ile Ile Ala Thr Asp Ile Gln Thr Lys Glu Leu Gln Lys Gln Ile Thr Lys Ile Gln Asn Phe Arg Val Tyr Tyr Arg Asp Ser Arg Asn Pro Leu Trp Lys Gly Pro Ala Lys Leu Leu Trp Lys Gly Glu Gly Ala Val Val Ile Gln Asp Asp Ser Asp Ile Lys Val Val Pro Arg Arg Lys Ala Lys Ile Ile Arg Asp Tyr Gly Lys Gln Met Ala Gly Asp Asp Cys Val Ala Ser Arg Gln Asp Glu Asp Glu Asp (SEQ ID NO:4).

10

15

20

25

30

As noted above, it will be understood that any combination of the mutations disclosed above may be suitable and therefore be utilized as an IA-pol-based adenoviral HIV vaccine of the present invention, either when administered alone or in a combined modality regime and/or a prime-boost regimen. For example, it may be possible to mutate only 2 of the 3 residues within the respective reverse transcriptase, RNase-H, and integrase coding regions while still abolishing these enzymatic activities. However, the IA-pol construct described above and disclosed as SEQ ID NO:3, as well as the expressed protein (SEQ ID NO:4;) is preferred. It is also preferred that at least one mutation be present in each of the three catalytic domains.

Another aspect of this portion of the invention are codon optimized HIV-1 Pol-based vaccine constructions which comprise a eukaryotic trafficking signal peptide such as from tPA (tissue-type plasminogen activator) or by a leader peptide such as is found in highly expressed mammalian proteins such as immunoglobulin leader peptides. Any functional leader peptide may be tested for efficacy. However, a preferred embodiment of the present invention, as with HIV-1 Nef constructs shown herein, is to provide for a HIV-1 Pol mutant adenoviral vaccine construction wherein the pol coding region or a portion thereof is operatively linked to a leader peptide, preferably a leader peptide from human tPA. In other words, a codon optimized HIV-1 Pol mutant such as IA-Pol (SEQ ID NO:4) may also comprise a leader peptide at the amino terminal portion of the protein, which may effect cellular trafficking and hence, immunogenicity of the expressed protein within the host cell. As noted in Figure 16A-B, a DNA vector which may be utilized to practice the present invention may be modified by known recombinant DNA methodology to contain a leader signal

peptide of interest, such that downstream cloning of the modified HIV-1 protein of interest results in a nucleotide sequence which encodes a modified HIV-1 tPA/Pol protein. In the alternative, as noted above, insertion of a nucleotide sequence which encodes a leader peptide may be inserted into a DNA vector housing the open reading frame for the Pol protein of interest. Regardless of the cloning strategy, the end result is a polynucleotide vaccine which comprises vector components for effective gene expression in conjunction with nucleotide sequences which encode a modified HIV-1 Pol protein of interest, including but not limited to a HIV-1 Pol protein which contains a leader peptide. The amino acid sequence of the human tPA leader utilized herein is as follows: MDAMKRGLCCVLLLCGAVFVSPSEISS (SEQ ID NO:17). Therefore, another aspect of the present invention is to generate HIV-1 Pol-based vaccine constructions which comprise a eukaryotic trafficking signal peptide such as from tPA. To this end, the present invention relates to a DNA molecule which encodes a codon optimized wt-pol DNA construct wherein the protease (PR) activity is deleted and a human tPA leader sequence is fused to the 5' end of the coding region. A DNA molecule which encodes this protein is disclosed herein as SEQ ID NO:5, the open reading frame disclosed herein as SEQ ID NO:6.

10

15

20

25

30

To this end, the present invention relates to a DNA molecule which encodes a codon optimized wt-pol DNA construct wherein the protease (PR) activity is deleted and a human tPA leader sequence is fused to the 5' end of the coding region (herein, "tPA-wt-pol"). A DNA molecule which encodes this protein is disclosed herein as SEQ ID NO:5, the open reading frame being contained from an initiating Met residue at nucleotides 8-10 to a termination codon from nucleotides 2633-2635. SEQ ID NO:5 is as follows:

GATCACCATG GATGCAATGA AGAGAGGGCT CTGCTGTGT CTGCTGTGT GTGGAGCAGT
CTTCGTTTCG CCCAGCGAGA TCTCCGCCCC CATCTCCCC ATTGAGACTG TGCCTGTGAA
GCTGAAGCCT GGCATGGATG GCCCCAAGGT GAAGCAGTGG CCCCTGACTG AGGAGAAGAT
CAAGGCCCTG GTGGAAATCT GCACTGAGAT GGAGAAGGAG GGCAAAATCT CCAAGATTGG
CCCCGAGAAC CCCTACAACA CCCCTGTGTT TGCCATCAAG AAGAAGGACT CCACCAAGTG
GAGGAAGCTG GTGGACTTCA GGGAGCTGAA CAAGAGGACC CAGGACTTCT GGGAGGTGCA
GCTGGGCATC CCCCACCCCG CTGGCCTGAA GAAGAAGAAG TCTGTGACTG TGCTGGATGT
GGGGGATGCC TACTTCTCTG TGCCCCTGGA TGAGGACTTC AGGAAGTACA CTGCCTTCAC
CATCCCCTCC ATCAACAATG AGACCCCTGG CATCAGGTAC CAGTACAATG TGCTGCCCCA
GGGCTGGAAG GGCTCCCCTG CCATCTTCCA GTCCTCCATG ACCAAGATCC TGGAGCCCTT
CAGGAAGCAG AACCCTGACA TTGTGATCTA CCAGTACATG GATGACCTGT ATGTGGGCTC
TGACCTGGAG ATTGGGCAGC ACAGGACCAA GATTGAGGAG CTGAGGCAGC ACCTGCTGAG

GTGGGGCCTG ACCACCCTG ACAAGAAGCA CCAGAAGGAG CCCCCCTTCC TGTGGATGGG CTATGAGCTG CACCCCGACA AGTGGACTGT GCAGCCCATT GTGCTGCCTG AGAAGGACTC CTGGACTGTG AATGACATCC AGAAGCTGGT GGGCAAGCTG AACTGGGCCT CCCAAATCTA CCCTGGCATC AAGGTGAGGC AGCTGTGCAA GCTGCTGAGG GGCACCAAGG CCCTGACTGA GGTGATCCCC CTGACTGAGG AGGCTGAGCT GGAGCTGGCT GAGAACAGGG AGATCCTGAA GGAGCCTGTG CATGGGGTGT ACTATGACCC CTCCAAGGAC CTGATTGCTG AGATCCAGAA GCAGGGCCAG GGCCAGTGGA CCTACCAAAT CTACCAGGAG CCCTTCAAGA ACCTGAAGAC TGGCAAGTAT GCCAGGATGA GGGGGGCCCA CACCAATGAT GTGAAGCAGC TGACTGAGGC TGTGCAGAAG ATCACCACTG AGTCCATTGT GATCTGGGGC AAGACCCCCA AGTTCAAGCT GCCCATCCAG AAGGAGACCT GGGAGACCTG GTGGACTGAG TACTGGCAGG CCACCTGGAT CCCTGAGTGG GAGTTTGTGA ACACCCCCC CCTGGTGAAG CTGTGGTACC AGCTGGAGAA GGAGCCCATT GTGGGGGCTG AGACCTTCTA TGTGGATGGG GCTGCCAACA GGGAGACCAA CCTGGCCAAG GCTGGCTATG TGACCAACAG GGGCAGGCAG AAGGTGGTGA CCCTGACTGA CACCACCAAC CAGAAGACTG AGCTCCAGGC CATCTACCTG GCCCTCCAGG ACTCTGGCCT 15 GGAGGTGAAC ATTGTGACTG ACTCCCAGTA TGCCCTGGGC ATCATCCAGG CCCAGCCTGA TCAGTCTGAG TCTGAGCTGG TGAACCAGAT CATTGAGCAG CTGATCAAGA AGGAGAAGGT GTACCTGGCC TGGGTGCCTG CCCACAAGGG CATTGGGGGC AATGAGCAGG TGGACAAGCT GGTGTCTGCT GGCATCAGGA AGGTGCTGTT CCTGGATGGC ATTGACAAGG CCCAGGATGA GCATGAGAAG TACCACTCCA ACTGGAGGGC TATGGCCTCT GACTTCAACC TGCCCCCTGT GGTGGCTAAG GAGATTGTGG CCTCCTGTGA CAAGTGCCAG CTGAAGGGGG AGGCCATGCA 20 TGGGCAGGTG GACTGCTCCC CTGGCATCTG GCAGCTGGAC TGCACCCACC TGGAGGGCAA GGTGATCCTG GTGGCTGTGC ATGTGGCCTC CGGCTACATT GAGGCTGAGG TGATCCCTGC TGAGACAGGC CAGGAGACTG CCTACTTCCT GCTGAAGCTG GCTGGCAGGT GGCCTGTGAA GACCATCCAC ACTGACAATG GCTCCAACTT CACTGGGGCC ACAGTGAGGG CTGCCTGCTG GTGGGCTGGC ATCAAGCAGG AGTTTGGCAT CCCCTACAAC CCCCAGTCCC AGGGGGTGGT 25 GGAGTCCATG AACAAGGAGC TGAAGAAGAT CATTGGGCAG GTGAGGGACC AGGCTGAGCA CCTGAAGACA GCTGTGCAGA TGGCTGTTGTT CATCCACAAC TTCAAGAGGA AGGGGGCCAT CGGGGGCTAC TCCGCTGGGG AGAGGATTGT GGACATCATT GCCACAGACA TCCAGACCAA GGAGCTCCAG AAGCAGATCA CCAAGATCCA GAACTTCAGG GTGTACTACA GGGACTCCAG GAACCCCTG TGGAAGGGCC CTGCCAAGCT GCTGTGGAAG GGGGAGGGGG CTGTGGTGAT 30 CCAGGACAAC TCTGACATCA AGGTGGTGCC CAGGAGGAAG GCCAAGATCA TCAGGGACTA TGGCAAGCAG ATGGCTGGGG ATGACTGTGT GGCCTCCAGG CAGGATGAGG ACTAAAGCCC GGGCAGATCT (SEQ ID NO:5).

The open reading frame of the wild type tPA-pol construct disclosed as SEQ ID NO:5 contains 875 amino acids, disclosed herein as SEQ ID NO:6, as follows:

Met Asp Ala Met Lys Arg Gly Leu Cys Cys Val Leu Leu Cys Gly

35

Ala Val Phe Val Ser Pro Ser Glu Ile Ser Ala Pro Ile Ser Pro Ile Glu Thr Val Pro Val Lys Leu Lys Pro Gly Met Asp Gly Pro Lys Val Lys Gln Trp Pro Leu Thr Glu Glu Lys Ile Lys Ala Leu Val Glu Ile Cys Thr Glu Met Glu Lys Glu Gly Lys Ile Ser Lys Ile Gly Pro Glu Asn Pro Tyr Asn Thr Pro Val Phe Ala Ile Lys Lys Lys Asp Ser Thr Lys Trp Arg Lys Leu Val Asp Phe Arg Glu Leu Asn Lys Arg Thr Gln Asp Phe Trp Glu Val Gln Leu Gly Ile Pro His Pro Ala Gly Leu Lys Lys Lys Ser Val Thr Val Leu Asp Val Gly Asp Ala Tyr Phe Ser Val Pro Leu Asp Glu Asp Phe Arg Lys Tyr Thr Ala Phe Thr Ile Pro 10 Ser Ile Asn Asn Glu Thr Pro Gly Ile Arg Tyr Gln Tyr Asn Val Leu Pro Gln Gly Trp Lys Gly Ser Pro Ala Ile Phe Gln Ser Ser Met Thr Lys Ile Leu Glu Pro Phe Arg Lys Gln Asn Pro Asp Ile Val Ile Tyr Gln Tyr Met Asp Asp Leu Tyr Val Gly Ser Asp Leu Glu Ile Gly Gln His Arg Thr Lys Ile Glu Glu Leu Arg Gln His Leu Leu Arg Trp Gly 15 Leu Thr Thr Pro Asp Lys Lys His Gln Lys Glu Pro Pro Phe Leu Trp Met Gly Tyr Glu Leu His Pro Asp Lys Trp Thr Val Gln Pro Ile Val Leu Pro Glu Lys Asp Ser Trp Thr Val Asn Asp Ile Gln Lys Leu Val Gly Lys Leu Asn Trp Ala Ser Gln Ile Tyr Pro Gly Ile Lys Val Arg Gln Leu Cys Lys Leu Leu Arg Gly Thr Lys Ala Leu Thr Glu Val Ile 20 Pro Leu Thr Glu Glu Ala Glu Leu Glu Leu Ala Glu Asn Arg Glu Ile Leu Lys Glu Pro Val His Gly Val Tyr Tyr Asp Pro Ser Lys Asp Leu Ile Ala Glu Ile Gln Lys Gln Gly Gln Gly Gln Trp Thr Tyr Gln Ile Tyr Gln Glu Pro Phe Lys Asn Leu Lys Thr Gly Lys Tyr Ala Arg Met Arg Gly Ala His Thr Asn Asp Val Lys Gln Leu Thr Glu Ala Val Gln 25 Lys Ile Thr Thr Glu Ser Ile Val Ile Trp Gly Lys Thr Pro Lys Phe Lys Leu Pro Ile Gln Lys Glu Thr Trp Glu Thr Trp Trp Thr Glu Tyr Trp Gln Ala Thr Trp Ile Pro Glu Trp Glu Phe Val Asn Thr Pro Pro Leu Val Lys Leu Trp Tyr Gln Leu Glu Lys Glu Pro Ile Val Gly Ala Glu Thr Phe Tyr Val Asp Gly Ala Ala Asn Arg Glu Thr Lys Leu Gly 30 Lys Ala Gly Tyr Val Thr Asn Arg Gly Arg Gln Lys Val Val Thr Leu Thr Asp Thr Thr Asn Gln Lys Thr.Glu Leu Gln Ala Ile Tyr Leu Ala Leu Gln Asp Ser Gly Leu Glu Val Asn Ile Val Thr Asp Ser Gln Tyr Ala Leu Gly Ile Ile Gln Ala Gln Pro Asp Gln Ser Glu Ser Glu Leu Val Asn Gln Ile Ile Glu Gln Leu Ile Lys Lys Glu Lys Val Tyr Leu 35 Ala Trp Val Pro Ala His Lys Gly Ile Gly Gly Asn Glu Gln Val Asp Lys Leu Val Ser Ala Gly Ile Arg Lys Val Leu Phe Leu Asp Gly Ile

Asp Lys Ala Gln Asp Glu His Glu Lys Tyr His Ser Asn Trp Arg Ala Met Ala Ser Asp Phe Asn Leu Pro Pro Val Val Ala Lys Glu Ile Val Ala Ser Cys Asp Lys Cys Gln Leu Lys Gly Glu Ala Met His Gly Gln Val Asp Cys Ser Pro Gly Ile Trp Gln Leu Asp Cys Thr His Leu Glu Gly Lys Val Ile Leu Val Ala Val His Val Ala Ser Gly Tyr Ile Glu 5 Ala Glu Val Ile Pro Ala Glu Thr Gly Gln Glu Thr Ala Tyr Phe Leu Leu Lys Leu Ala Gly Arg Trp Pro Val Lys Thr Ile His Thr Asp Asn Gly Ser Asn Phe Thr Gly Ala Thr Val Arg Ala Ala Cys Trp Trp Ala Gly Ile Lys Gln Glu Phe Gly Ile Pro Tyr Asn Pro Gln Ser Gln Gly Val Val Glu Ser Met Asn Lys Glu Leu Lys Lys Ile Ile Gly Gln Val 10 Arg Asp Gln Ala Glu His Leu Lys Thr Ala Val Gln Met Ala Val Phe Ile His Asn Phe Lys Arg Lys Gly Gly Ile Gly Gly Tyr Ser Ala Gly Glu Arg Ile Val Asp Ile Ile Ala Thr Asp Ile Gln Thr Lys Glu Leu Gln Lys Gln Ile Thr Lys Ile Gln Asn Phe Arg Val Tyr Tyr Arg Asp 15 Ser Arg Asn Pro Leu Trp Lys Gly Pro Ala Lys Leu Leu Trp Lys Gly Glu Gly Ala Val Val Ile Gln Asp Asn Ser Asp Ile Lys Val Val Pro Arg Arg Lys Ala Lys Ile Ile Arg Asp Tyr Gly Lys Gln Met Ala Gly Asp Asp Cys Val Ala Ser Arg Gln Asp Glu Asp (SEQ ID NO:6).

20

25

30

The present invention also relates to a codon optimized HIV-1 Pol mutant contained within a recombinant adenoviral vector such as IA-Pol (SEQ ID NO:4) which comprises a leader peptide at the amino terminal portion of the protein, which may effect cellular trafficking and hence, immunogenicity of the expressed protein within the host cell. Any such adenoviral-based HIV-1 DNA pol mutant disclosed in the above paragraphs is suitable for fusion downstream of a leader peptide, such as a leader peptide including but not limited to the human tPA leader sequence. Therefore, any such leader peptide-based HIV-1 pol mutant construct may include but is not limited to a mutated DNA molecule which effectively alters the catalytic activity of the RT, RNase and/or IN region of the expressed protein, resulting in at least substantially decreased enzymatic activity one or more of the RT, RNase H and/or IN functions of HIV-1 Pol. In a preferred embodiment of this portion of the invention, a leader peptide/HIV-1 DNA pol construct contains a mutation or mutations within the Pol coding region which effectively abolishes RT, RNase H and IN activity. An especially preferable HIV-1 DNA pol construct is a DNA molecule which contains at least one point mutation which alters the active site and catalytic activity within the RT, RNase H and IN domains of Pol, such that each activity is at least substantially abolished, and preferably totally abolished. Such a HIV-1 Pol mutant will most likely

comprise at least one point mutation in or around each catalytic domain responsible for RT, RNase H and IN activity, respectfully. An especially preferred embodiment of this portion of the invention relates to a human tPA leader fused to the IA-Pol protein comprising the nine mutations shown in Table 1. The DNA molecule is disclosed herein as SEQ ID NO:7 and the expressed tPA-IA Pol protein comprises a fusion 5 junction as shown in Figure 18. The complete amino acid sequence of the expressed protein is set forth in SEQ ID NO:8. To this end, SEQ ID NO:7 discloses the nucleotide sequence which codes for a human tPA leader fused to the IA Pol protein comprising the nine mutations shown in Table 1 (herein, "tPA-opt-IApol"). The open reading frame begins with the initiating Met (nucleotides 8-10) and terminates with a "TAA" codon at nucleotides 2633-2635. The nucleotide sequence encoding tPA-IAPol is also disclosed as follows: GATCACCATG GATGCAATGA AGAGAGGGCT CTGCTGTGTG CTGCTGCTGT GTGGAGCAGT CTTCGTTTCG CCCAGCGAGA TCTCCGCCCC CATCTCCCCC ATTGAGACTG TGCCTGTGAA GCTGAAGCCT GGCATGGATG GCCCCAAGGT GAAGCAGTGG CCCCTGACTG AGGAGAAGAT CAAGGCCCTG GTGGAAATCT GCACTGAGAT GGAGAAGGAG GGCAAAATCT CCAAGATTGG CCCCGAGAAC CCCTACAACA CCCCTGTGTT TGCCATCAAG AAGAAGGACT CCACCAAGTG GAGGAAGCTG GTGGACTTCA GGGAGCTGAA CAAGAGGACC CAGGACTTCT GGGAGGTGCA GCTGGGCATC CCCCACCCCG CTGGCCTGAA GAAGAAGAAG TCTGTGACTG TGCTGGCTGT GGGGGATGCC TACTTCTCTG TGCCCCTGGA TGAGGACTTC AGGAAGTACA CTGCCTTCAC CATCCCCTCC ATCAACAATG AGACCCCTGG CATCAGGTAC CAGTACAATG TGCTGCCCCA GGGCTGGAAG GGCTCCCCTG CCATCTTCCA GTCCTCCATG ACCAAGATCC TGGAGCCCTT CAGGAAGCAG AACCCTGACA TTGTGATCTA CCAGTACATG GCTGCCCTGT ATGTGGGCTC TGACCTGGAG ATTGGGCAGC ACAGGACCAA GATTGAGGAG CTGAGGCAGC ACCTGCTGAG GTGGGGCCTG ACCACCCCTG ACAAGAAGCA CCAGAAGGAG CCCCCCTTCC TGTGGATGGG CTATGAGCTG CACCCCGACA AGTGGACTGT GCAGCCCATT GTGCTGCCTG AGAAGGACTC CTGGACTGTG AATGACATCC AGAAGCTGGT GGGCAAGCTG AACTGGGCCT CCCAAATCTA CCCTGGCATC AAGGTGAGGC AGCTGTGCAA GCTGCTGAGG GGCACCAAGG CCCTGACTGA GGTGATCCCC CTGACTGAGG AGGCTGAGCT GGAGCTGGCT GAGAACAGGG AGATCCTGAA GGAGCCTGTG CATGGGGTGT ACTATGACCC CTCCAAGGAC CTGATTGCTG AGATCCAGAA -GCAGGGCCAG GGCCAGTGGA CCTACCAAAT CTACCAGGAG CCCTTCAAGA ACCTGAAGAC TGGCAAGTAT GCCAGGATGA GGGGGGCCCA CACCAATGAT GTGAAGCAGC TGACTGAGGC TGTGCAGAAG ATCACCACTG AGTCCATTGT GATCTGGGGC AAGACCCCCA AGTTCAAGCT GCCCATCCAG AAGGAGACCT GGGAGACCTG GTGGACTGAG TACTGGCAGG CCACCTGGAT CCCTGAGTGG GAGTTTGTGA ACACCCCCCC CCTGGTGAAG CTGTGGTACC AGCTGGAGAA GGAGCCCATT GTGGGGGCTG AGACCTTCTA TGTGGCTGGG GCTGCCAACA GGGAGACCAA

10

15

20

25

30

35

GCTGGGCAAG GCTGGCTATG TGACCAACAG GGGCAGGCAG AAGGTGGTGA CCCTGACTGA CACCACCAAC CAGAAGACTG CCCTCCAGGC CATCTACCTG GCCCTCCAGG ACTCTGGCCT GGAGGTGAAC ATTGTGACTG CCTCCCAGTA TGCCCTGGGC ATCATCCAGG CCCAGCCTGA TCAGTCTGAG TCTGAGCTGG TGAACCAGAT CATTGAGCAG CTGATCAAGA AGGAGAAGGT GTACCTGGCC TGGGTGCCTG CCCACAAGGG CATTGGGGGC AATGAGCAGG TGGACAAGCT GGTGTCTGCT GGCATCAGGA AGGTGCTGTT CCTGGATGGC ATTGACAAGG CCCAGGATGA GCATGAGAAG TACCACTCCA ACTGGAGGGC TATGGCCTCT GACTTCAACC TGCCCCCTGT GGTGGCTAAG GAGATTGTGG CCTCCTGTGA CAAGTGCCAG CTGAAGGGGG AGGCCATGCA TGGGCAGGTG GACTGCTCCC CTGGCATCTG GCAGCTGGCC TGCACCCACC TGGAGGGCAA 10 GGTGATCCTG GTGGCTGTGC ATGTGGCCTC CGGCTACATT GAGGCTGAGG TGATCCCTGC TGAGACAGGC CAGGAGACTG CCTACTTCCT GCTGAAGCTG GCTGGCAGGT GGCCTGTGAA GACCATCCAC ACTGCCAATG GCTCCAACTT CACTGGGGCC ACAGTGAGGG CTGCCTGCTG GTGGGCTGGC ATCAAGCAGG AGTTTGGCAT CCCCTACAAC CCCCAGTCCC AGGGGGTGGT GGCCTCCATG AACAAGGAGC TGAAGAAGAT CATTGGGCAG GTGAGGGACC AGGCTGAGCA CCTGAAGACA GCTGTGCAGA TGGCTGTGTT CATCCACAAC TTCAAGAGGA AGGGGGGCAT 15 CGGGGGCTAC TCCGCTGGGG AGAGGATTGT GGACATCATT GCCACAGACA TCCAGACCAA GGAGCTCCAG AAGCAGATCA CCAAGATCCA GAACTTCAGG GTGTACTACA GGGACTCCAG GAACCCCTG TGGAAGGGCC CTGCCAAGCT GCTGTGGAAG GGGGAGGGGG CTGTGGTGAT CCAGGACAAC TCTGACATCA AGGTGGTGCC CAGGAGGAAG GCCAAGATCA TCAGGGACTA TGGCAAGCAG ATGGCTGGGG ATGACTGTGT GGCCTCCAGG CAGGATGAGG ACTAAAGCCC 20 GGGCAGATCT (SEQ ID NO:7).

The open reading frame of the tPA-IA-pol construct disclosed as SEQ ID NO:7 contains 875 amino acids, disclosed herein as tPA-IA-Pol and SEQ ID NO:8, as follows:

25 Met Asp Ala Met Lys Arg Gly Leu Cys Cys Val Leu Leu Leu Cys Gly Ala Val Phe Val Ser Pro Ser Glu Ile Ser Ala Pro Ile Ser Pro Ile Glu Thr Val Pro Val Lys Leu Lys Pro Gly Met Asp Gly Pro Lys Val Lys Gln Trp Pro Leu Thr Glu Glu Lys Ile Lys Ala Leu Val Glu Ile Cys Thr Glu Met Glu Lys Glu Gly Lys Ile Ser Lys Ile Gly Pro Glu Asn Pro Tyr Asn Thr Pro Val Phe Ala Ile Lys Lys Lys Asp Ser Thr Lys Trp Arg Lys Leu Val Asp Phe Arg Glu Leu Asn Lys Arg Thr Gln Asp Phe Trp Glu Val Gln Leu Gly Ile Pro His Pro Ala Gly Leu Lys Lys Lys Lys Ser Val Thr Val Leu Ala Val Gly Asp Ala Tyr Phe Ser Val Pro Leu Asp Glu Asp Phe Arg Lys Tyr Thr Ala Phe Thr Ile Pro Ser Ile Asn Asn Glu Thr Pro Gly Ile Arg Tyr Gln Tyr Asn Val Leu Pro Gln Gly Trp Lys Gly Ser Pro Ala Ile Phe Gln Ser Ser Met Thr

Lys Ile Leu Glu Pro Phe Arg Lys Gln Asn Pro Asp Ile Val Ile Tyr Gln Tyr Met Ala Ala Leu Tyr Val Gly Ser Asp Leu Glu Ile Gly Gln His Arg Thr Lys Ile Glu Glu Leu Arg Gln His Leu Leu Arg Trp Gly Leu Thr Thr Pro Asp Lys Lys His Gln Lys Glu Pro Pro Phe Leu Trp Met Gly Tyr Glu Leu His Pro Asp Lys Trp Thr Val Gln Pro Ile Val Leu Pro Glu Lys Asp Ser Trp Thr Val Asn Asp Ile Gln Lys Leu Val Gly Lys Leu Asn Trp Ala Ser Gln Ile Tyr Pro Gly Ile Lys Val Arg Gln Leu Cys Lys Leu Leu Arg Gly Thr Lys Ala Leu Thr Glu Val Ile Pro Leu Thr Glu Glu Ala Glu Leu Glu Leu Ala Glu Asn Arg Glu Ile 10 Leu Lys Glu Pro Val His Gly Val Tyr Tyr Asp Pro Ser Lys Asp Leu Ile Ala Glu Ile Gln Lys Gln Gly Gln Gly Gln Trp Thr Tyr Gln Ile Tyr Gln Glu Pro Phe Lys Asn Leu Lys Thr Gly Lys Tyr Ala Arg Met Arg Gly Ala His Thr Asn Asp Val Lys Gln Leu Thr Glu Ala Val Gln Lys Ile Thr Thr Glu Ser Ile Val Ile Trp Gly Lys Thr Pro Lys Phe 15 Lys Leu Pro Ile Gln Lys Glu Thr Trp Glu Thr Trp Trp Thr Glu Tyr Trp Gln Ala Thr Trp Ile Pro Glu Trp Glu Phe Val Asn Thr Pro Pro Leu Val Lys Leu Trp Tyr Gln Leu Glu Lys Glu Pro Ile Val Gly Ala Glu Thr Phe Tyr Val Ala Gly Ala Ala Asn Arg Glu Thr Lys Leu Gly Lys Ala Gly Tyr Val Thr Asn Arg Gly Arg Gln Lys Val Val Thr Leu 20 Thr Asp Thr Thr Asn Gln Lys Thr Ala Leu Gln Ala Ile Tyr Leu Ala Leu Gln Asp Ser Gly Leu Glu Val Asn Ile Val Thr Ala Ser Gln Tyr Ala Leu Gly Ile Ile Gln Ala Gln Pro Asp Gln Ser Glu Ser Glu Leu Val Asn Gln Ile Ile Glu Gln Leu Ile Lys Lys Glu Lys Val Tyr Leu Ala Trp Val Pro Ala His Lys Gly Ile Gly Gly Asn Glu Gln Val Asp Lys Leu Val Ser Ala Gly Ile Arg Lys Val Leu Phe Leu Asp Gly Ile 25 Asp Lys Ala Gln Asp Glu His Glu Lys Tyr His Ser Asn Trp Arg Ala Met Ala Ser Asp Phe Asn Leu Pro Pro Val Val Ala Lys Glu Ile Val Ala Ser Cys Asp Lys Cys Gln Leu Lys Gly Glu Ala Met His Gly Gln Val Asp Cys Ser Pro Gly Ile Trp Gln Leu Ala Cys Thr His Leu Glu Gly Lys Val Ile Leu Val Ala Val His Val Ala Ser Gly Tyr Ile Glu 30 Ala Glu Val Ile Pro Ala Glu Thr Gly Gln Glu Thr Ala Tyr Phe Leu Leu Lys Leu Ala Gly Arg Trp Pro Val Lys Thr Ile His Thr Ala Asn Gly Ser Asn Phe Thr Gly Ala Thr Val Arg Ala Ala Cys Trp Trp Ala Gly Ile Lys Gln Glu Phe Gly Ile Pro Tyr Asn Pro Gln Ser Gln Gly 35 Val Val Ala Ser Met Asn Lys Glu Leu Lys Lys Ile Ile Gly Gln Val Arg Asp Gln Ala Glu His Leu Lys Thr Ala Val Gln Met Ala Val Phe

Ile His Asn Phe Lys Arg Lys Gly Gly Ile Gly Gly Tyr Ser Ala Gly Glu Arg Ile Val Asp Ile Ile Ala Thr Asp Ile Gln Thr Lys Glu Leu Gln Lys Gln Ile Thr Lys Ile Gln Asn Phe Arg Val Tyr Tyr Arg Asp Ser Arg Asn Pro Leu Trp Lys Gly Pro Ala Lys Leu Leu Trp Lys Gly Glu Gly Ala Val Val Ile Gln Asp Asn Ser Asp Ile Lys Val Val Pro Arg Arg Lys Ala Lys Ile Ile Arg Asp Tyr Gly Lys Gln Met Ala Gly Asp Asp Cys Val Ala Ser Arg Gln Asp Glu Asp (SEQ ID NO:8).

#### **EXAMPLE 18**

# CODON OPTIMIZED HIV-1 NEF AND CODON OPTIMIZED HIV-1 NEF MODIFICATIONS

10

15

20

25

30

Codon optimized version of HIV-1 Nef and HIV-1 Nef modifications are essentially as described in U.S. Application Serial No. 09/738,782, filed December 15, 2000 and PCT International Application PCT/US00/34162, also filed December 15, 2000, both documents which are hereby incorporated by reference. As disclosed within the above-mentioned documents, particular embodiments of codon optimized Nef and Nef modifications relate to a DNA molecule encoding HIV-1 Nef from the HIV-1 ifrl isolate wherein the codons are optimized for expression in a mammalian system such as a human. The DNA molecule which encodes this protein is disclosed herein as SEQ ID NO:9, while the expressed open reading frame is disclosed herein as SEQ ID NO:10. Another embodiment of Nef-based coding regions for use in the adenoviral vectors of the present invention comprise a codon optimized DNA molecule encoding a protein containing the human plasminogen activator (tpa) leader peptide fused with the NH2-terminus of the HIV-1 Nef polypeptide. The DNA molecule which encodes this protein is disclosed herein as SEO ID NO:11, while the expressed open reading frame is disclosed herein as SEO ID NO:12. Another modified Nef optimized coding region relates to a DNA molecule encoding optimized HIV-1 Nef wherein the open reading frame codes for modifications at the amino terminal myristylation site (Gly-2 to Ala-2) and substitution of the Leu-174-Leu-175 dileucine motif to Ala-174-Ala-175, herein described as opt nef (G2A, LLAA). The DNA molecule which encodes this protein is disclosed herein as SEQ ID NO:13, while the expressed open reading frame is disclosed herein as SEQ ID NO:14. An additional embodiment relates to a DNA molecule encoding optimized HIV-1 Nef wherein the amino terminal myristylation site and dileucine motif have been deleted, as well as comprising a tPA leader peptide. This DNA molecule, opt tpanef (LLAA), comprises an open reading frame which

encodes a Nef protein containing a tPA leader sequence fused to amino acid residue 6-216 of HIV-1 Nef (jfrl), wherein Leu-174 and Leu-175 are substituted with Ala-174 and Ala-175, herein referred to as opt tpanef (LLAA) is disclosed herein as SEQ ID NO:15, while the expressed open reading frame is disclosed herein as SEO ID NO:16.

As disclosed in the above-identified documents (U.S. Application Serial No. 09/738,782 and PCT International Application PCT/US00/34162) and reiterated herein, the following nef-based nucleotide and amino acid sequences which comprise the respective open reading frame are as follows:

5

10

15

. 20

25

30

The nucleotide sequence of the codon optimized version of HIV-1 jrfl nef gene is disclosed herein as SEQ ID NO:9, as shown herein:

GATCTGCCAC CATGGGCGGC AAGTGGTCCA AGAGGTCCGT GCCCGGCTGG TCCACCGTGA GGGAGAGGAT GAGGAGGGCC GAGCCCGCCG CCGACAGGGT GAGGAGGACC GAGCCCGCCG CCGTGGCGT GGGCGCCGTG TCCAGGGACC TGGAGAAGCA CGGCGCCATC ACCTCCTCCA ACACCGCCGC CACCAACGCC GACTGCGCCT GGCTGGAGGC CCAGGAGGAC GAGGAGGTGG GCTTCCCCGT GAGGCCCCAG GTGCCCCTGA GGCCCATGAC CTACAAGGGC GCCGTGGACC TGTCCCACTT CCTGAAGGAG AAGGGCGGCC TGGAGGGCCT GATCCACTCC CAGAAGAGGC AGGACATCCT GGACCTGTGG GTGTACCACA CCCAGGGCTA CTTCCCCGAC TGGCAGAACT ACACCCCGG CCCCGGCATC AGGTTCCCCC TGACCTTCGG CTGGTGCTTC AAGCTGGTGC CCGTGGAGCC CGAGAAGGTG GAGGAGGCCA ACGAGGGCGA GAACAACTGC CTGCTGCACC CCATGTCCCA GCACGGCATC GAGGACCCCG AGAAGGAGGT GCTGGAGTGG AGGTTCGACT CCAAGCTGGC CTTCCACCAC GTGGCCAGGG AGCTGCACCC CGAGTACTAC AAGGACTGCT AAAGCCCGGG C (SEQ ID NO:9).

Preferred codon usage is as follows: Met (ATG), Gly (GGC), Lys (AAG), Trp (TGG), Ser (TCC), Arg (AGG), Val (GTG), Pro (CCC), Thr (ACC), Glu (GAG); Leu (CTG), His (CAC), Ile (ATC), Asn (AAC), Cys (TGC), Ala (GCC), Gln (CAG), Phe (TTC) and Tyr (TAC). For an additional discussion relating to mammalian (human) codon optimization, see WO 97/31115 (PCT/US97/02294), which is hereby incorporated by reference. See also Figure 19A-B for a comparion of wild type vs. codon optimized nucleotides comprising the open reading frame of HIV-Nef.

The open reading frame for SEQ ID NO:9 above comprises an initiating methionine residue at nucleotides 12-14 and a "TAA" stop codon from nucleotides 660-662. The open reading frame of SEQ ID NO:9 provides for a 216 amino acid HIV-1 Nef protein expressed through utilization of a codon optimized DNA vaccine vector. The 216 amino acid HIV-1 Nef (ifrl) protein is disclosed herein as SEQ ID 35 NO:10, and as follows:

Met Gly Gly Lys Trp Ser Lys Arg Ser Val Pro Gly Trp Ser Thr Val

Arg Glu Arg Met Arg Ala Glu Pro Ala Ala Asp Arg Val Arg Arg Arg Clu Pro Ala Ala Asp Arg Val Arg Arg Thr Glu Pro Ala Ala Ala Val Gly Val Gly Ala Val Ser Arg Asp Leu Glu Lys His Gly Ala Ile Thr Ser Ser Asn Thr Ala Ala Thr Asn Ala Asp Cys Ala Trp Leu Glu Ala Gln Glu Asp Glu Glu Val Gly Phe Pro Val Arg Pro Gln Val Pro Leu Arg Pro Met Thr Tyr Lys Gly Ala Val Asp Leu Ser His Phe Leu Lys Glu Lys Gly Gly Leu Glu Gly Leu Ile His Ser Gln Lys Arg Gln Asp Ile Leu Asp Leu Trp Val Tyr His Thr Gln Gly Tyr Phe Pro Asp Trp Gln Asn Tyr Thr Pro Gly Pro Gly Ile Arg Phe Pro Leu Thr Phe Gly Trp Cys Phe Lys Leu Val Pro Val Glu Pro Glu Lys Val Glu Glu Ala Asn Glu Gly Glu Asn Asn Cys Leu Leu His Pro Met Ser Gln His Gly Ile Glu Asp Pro Glu Lys Glu Val Leu Glu His Pro Gly Pro Glu Tyr Tyr Lys Asp Cys (SEQ ID NO:10).

10

15

20

25

30

35

HIV-1 Nef is a 216 amino acid cytosolic protein which associates with the inner surface of the host cell plasma membrane through myristylation of Gly-2 (Franchini et al., 1986, Virology 155: 593-599). While not all possible Nef functions have been elucidated, it has become clear that correct trafficking of Nef to the inner plasma membrane promotes viral replication by altering the host intracellular environment to facilitate the early phase of the HIV-1 life cycle and by increasing the infectivity of progeny viral particles. In one aspect of the invention regarding codon-optimized, protein-modified polypeptides, the nef-encoding region of the adenovirus vector of the present invention is modified to contain a nucleotide sequence which encodes a heterologous leader peptide such that the amino terminal region of the expressed protein will contain the leader peptide. The diversity of function that typifies eukaryotic cells depends upon the structural differentiation of their membrane boundaries. To generate and maintain these structures, proteins must be transported from their site of synthesis in the endoplasmic reticulum to predetermined destinations throughout the cell. This requires that the trafficking proteins display sorting signals that are recognized by the molecular machinery responsible for route selection located at the access points to the main trafficking pathways. Sorting decisions for most proteins need to be made only once as they traverse their biosynthetic pathways since their final destination, the cellular location at which they perform their function, becomes their permanent residence. Maintenance of intracellular integrity depends in part on the selective sorting and accurate transport of proteins to their correct destinations. Defined sequence motifs exist in proteins which can act as 'address labels'. A number of sorting signals have

been found associated with the cytoplasmic domains of membrane proteins. An effective induction of CTL responses often required sustained, high level endogenous expression of an antigen. As membrane-association via myristylation is an essential requirement for most of Nef's function, mutants lacking myristylation, by glycine-to-alanine change, change of the dileucine motif and/or by substitution with a tpa leader sequence as described herein, will be functionally defective, and therefore will have improved safety profile compared to wild-type Nef for use as an HIV-1 vaccine component.

In another embodiment of this portion of the invention, either the DNA vector or the HIV-1 nef nucleotide sequence is modified to include the human tissue-specific plasminogen activator (tPA) leader. As shown in Figure 16A-B, a DNA vector may be modified by known recombinant DNA methodology to contain a leader signal peptide of interest, such that downstream cloning of the modified HIV-1 protein of interest results in a nucleotide sequence which encodes a modified HIV-1 tPA/Nef protein. In the alternative, as noted above, insertion of a nucleotide sequence which encodes a leader peptide may be inserted into a DNA vector housing the open reading frame for the Nef protein of interest. Regardless of the cloning strategy, the end result is a polynucleotide vaccine which comprises vector components for effective gene expression in conjunction with nucleotide sequences which encode a modified HIV-1 Nef protein of interest, including but not limited to a HIV-1 Nef protein which contains a leader peptide. The amino acid sequence of the human tPA leader utilized herein is as follows: MDAMKRGLCCVLLLCGAVFVSPSEISS (SEQ ID NO:17).

It has been shown that myristylation of Gly-2 in conjunction with a dileucine motif in the carboxy region of the protein is essential for Nef-induced down regulation of CD4 (Aiken et al., 1994, Cell 76: 853-864) via endocytosis. It has also been shown that Nef expression promotes down regulation of MHCI (Schwartz et al., 1996, Nature Medicine 2(3): 338-342) via endocytosis. The present invention relates in part to DNA vaccines which encode modified Nef proteins altered in trafficking and/or functional properties. The modifications introduced into the adenoviral vector HIV vaccines of the present invention include but are not limited to additions, deletions or substitutions to the nef open reading frame which results in the expression of a modified Nef protein which includes an amino terminal leader peptide, modification or deletion of the amino terminal myristylation site, and modification or deletion of the dileucine motif within the Nef protein and which alter function within the infected host cell. Therefore, a central theme of the DNA molecules and recombinant adenoviral HIV vaccines of the present invention is (1)

host administration and intracellular delivery of a codon optimized nef-based adenoviral HIV vaccine; (2) expression of a modified Nef protein which is immunogenic in terms of eliciting both CTL and Th responses; and, (3) inhibiting or at least altering known early viral functions of Nef which have been shown to promote HIV-1 replication and load within an infected host. Therefore, the nef coding region may be altered, resulting in a DNA vaccine which expresses a modified Nef protein wherein the amino terminal Gly-2 myristylation residue is either deleted or modified to express alternate amino acid residues. Also, the nef coding region may be altered so as to result in a DNA vaccine which expresses a modified Nef protein wherein the dileucine motif is either deleted or modified to express alternate amino acid residues. In addition, the adenoviral vector HIV vaccines of the present invention also relate to an isolated DNA molecule, regardless of codon usage, which expresses a wild type or modified Nef protein as described herein, including but not limited to modified Nef proteins which comprise a deletion or substitution of Gly 2, a deletion or substitution of Leu 174 and Leu 175 and/or inclusion of a leader sequence.

10

15

20

25

30

35

Therefore, specific Nef-based constructs further include the following, as exemplification's and not limitations. For example, the present invention relates to an adenoviral vector vaccine which encodes modified forms of HIV-1, an open reading frame which encodes a Nef protein which comprises a tPA leader sequence fused to amino acid residue 6-216 of HIV-1 Nef (jfrl) is referred to herein as opt tpanef. The nucleotide sequence comprising the open reading frame of opt tpanef is disclosed herein as SEQ ID NO:11, as shown below:

CATGGATGCA ATGAAGAGA GGCTCTGCTG TGTGCTGCTG CTGTGTGGAG CAGTCTTCGT
TTCGCCCAGC GAGATCTCCT CCAAGAGGTC CGTGCCCGGC TGGTCCACCG TGAGGGAGAG
GATGAGGAGG GCCGAGCCCG CCGCCGACAG GGTGAGGAGG ACCGAGCCCG CCGCCGTGGG
CGTGGGCGCC GTGTCCAGGG ACCTGGAGAA GCACGGCGCC ATCACCTCCT CCAACACCGC
CGCCACCAAC GCCGACTGCG CCTGGCTGGA GGCCCAGGAG GACGAGGAGG TGGGCTTCCC
CGTGAGGCCC CAGGTGCCCC TGAGGCCCAT GACCTACAAG GGCGCCGTGG ACCTGTCCCA
CTTCCTGAAG GAGAAGGGCG GCCTGGAGGG CCTGATCCAC TCCCAGAAGA GGCAGGACAT
CCTGGACCTG TGGGTGTACC ACACCCAGGG CTACTTCCCC GACTGGCAGA ACTACACCCC
CGGCCCCGGC ATCAGGTTCC CCCTGACCTT CGGCTGGTGC TTCAAGCTGG TGCCCGTGGA
GCCCGAGAAG GTGGAGGAGG CCAACGAGGG CGAGAACAAC TGCCTGCTGC ACCCCATGTC
CCAGCACGGC ATCGAGGACC CCGAGAAGGA GGTCCTGGAG TGGAGGTTCG ACCCCATGTC
GGCCTTCCAC CACGTGGCCA GGGAGCTGCA CCCCGAGTAC TACAAGGACT GCTAAAGCC
(SEO ID NO:11).

The open reading frame for SEQ ID NO:11 comprises an initiating methionine

residue at nucleotides 2-4 and a "TAA" stop codon from nucleotides 713-715. The open reading frame of SEQ ID NO:3 provides for a 237 amino acid HIV-1 Nef protein which comprises a tPA leader sequence fused to amino acids 6-216 of HIV-1 Nef, including the dileucine motif at amino acid residues 174 and 175. This 237 amino acid tPA/Nef (jfrl) fusion protein is disclosed herein as SEQ ID NO:12, and is shown as follows:

Met Asp Ala Met Lys Arg Cly Leu Cys Cys Val Leu Leu Leu Cys Cly Ala Val Phe Val Ser Pro Ser Glu Ile Ser Ser Lys Arg Ser Val Pro Gly Trp Ser Thr Val Arg Glu Arg Met Arg Arg Ala Glu Pro Ala Ala 10 Asp Arg Val Arg Arg Thr Glu Pro Ala Ala Val Gly Val Gly Ala Val Ser Arg Asp Leu Glu Lys His Gly Ala Ile Thr Ser Ser Asn Thr Ala Ala Thr Asn Ala Asp Cys Ala Trp Leu Glu Ala Gln Glu Asp Glu Glu Val Gly Phe Pro Val Arg Pro Gln Val Pro Leu Arg Pro Met Thr Tyr Lys Gly Ala Val Asp Leu Ser His Phe Leu Lys Glu Lys Gly Gly Leu Glu Gly Leu Ile His Ser Gln Lys Arg Gln Asp Ile Leu Asp Leu Trp 15 Val Tyr His Thr Gln Gly Tyr Phe Pro Asp Trp Gln Asn Tyr Thr Pro Gly Pro Gly Ile Arg Phe Pro Leu Thr Phe Gly Trp Cys Phe Lys Leu Val Pro Val Glu Pro Glu Lys Val Glu Glu Ala Asn Glu Gly Glu Asn Asn Cys Leu Leu His Pro Met Ser Gln His Gly Ile Glu Asp Pro Glu 20 Lys Glu Val Leu Glu Trp Arg Phe Asp Ser Lys Leu Ala Phe His His Val Ala Arg Glu Leu His Pro Glu Tyr Tyr Lys Asp Cys (SEQ ID NO:12). Therefore, this exemplified Nef protein, Opt tPA-Nef, contains both a tPA leader sequence as well as deleting the myristylation site of Gly-2A DNA molecule encoding HIV-1 Nef from the HIV-1 ifrl isolate wherein the codons are optimized for expression in a mammalian system such as a human. 25

In another specific embodiment of the present invention, a DNA molecule is disclosed which encodes optimized HIV-1 Nef wherein the open reading frame of a recombinant adenoviral HIV vaccine encodes for modifications at the amino terminal myristylation site (Gly-2 to Ala-2) and substitution of the Leu-174-Leu-175 dileucine motif to Ala-174-Ala-175. This open reading frame is herein described as opt nef (G2A,LLAA) and is disclosed as SEQ ID NO:13, which comprises an initiating methionine residue at nucleotides 12-14 and a "TAA" stop codon from nucleotides 660-662. The nucleotide sequence of this codon optimized version of HIV-1 jrfl nef gene with the above mentioned modifications is disclosed herein as SEQ ID NO:13, as follows:

35 as follows:

30

GATCTGCCAC CATGGCCGGC AAGTGGTCCA AGAGGTCCGT GCCCGGCTGG TCCACCGTGA
GGGAGAGGAT GAGGAGGGCC GAGCCCGCCG CCGACAGGGT GAGGAGGACC GAGCCCGCCG
CCGTGGGCGT GGGCGCCGTG TCCAGGGACC TGGAGAAGCA CGGCGCCATC ACCTCCTCCA
ACACCGCCGC CACCAACGCC GACTGCGCCT GGCTGGAGGC CCAGGAGGAC GAGGAGGTGG
GCTTCCCCGT GAGGCCCCAG GTGCCCCTGA GGCCCATGAC CTACAAGGGC GCCGTGGACC
TGTCCCACTT CCTGAAGGAG AAGGGCGGCC TGGAGGGCCT GATCCACTCC CAGAAGAGGC
AGGACATCCT GGACCTGTGG GTGTACCACA CCCAGGGCTA CTTCCCCGAC TGGCAGAACT
ACACCCCCGG CCCCGGCATC AGGTTCCCCC TGACCTTCGG CTGGTGCTTC AAGCTGGTGC
CCGTGGAGCC CGAGAAGGTG GAGGAGGCCA ACGAGGGCGA GAACAACTGC GCCGCCCACC
CCATGTCCCA GCACGGCATC GAGGACCCCG AGAAGGAGGT GCTGGAGTGG AGGTTCGACT
CCAAGCTGGC CTTCCACCAC GTGGCCAGGG AGCTGCACCC CGAGTACTAC AAGGACTGCT
AAAGCCCGGG C (SEQ ID NO:13).

The open reading frame of SEQ ID NO:13 encodes Nef (G2A,LLAA), disclosed herein as SEQ ID NO:14, as follows:

10

30

35

Met Ala Gly Lys Trp Ser Lys Arg Ser Val Pro Gly Trp Ser Thr Val 15 Arg Glu Arg Met Arg Arg Ala Glu Pro Ala Ala Asp Arg Val Arg Arg Thr Glu Pro Ala Ala Val Gly Val Gly Ala Val Ser Arg Asp Leu Glu Lys His Gly Ala Ile Thr Ser Ser Asn Thr Ala Ala Thr Asn Ala Asp Cys Ala Trp Leu Glu Ala Gln Glu Asp Glu Glu Val Gly Phe Pro Val Arg Pro Gln Val Pro Leu Arg Pro Met Thr Tyr Lys Gly Ala Val Asp 20 Leu Ser His Phe Leu Lys Glu Lys Gly Gly Leu Glu Gly Leu Ile His Ser Gln Lys Arg Gln Asp Ile Leu Asp Leu Trp Val Tyr His Thr Gln Gly Tyr Phe Pro Asp Trp Gln Asn Tyr Thr Pro Gly Pro Gly Ile Arg Phe Pro Leu Thr Phe Gly Trp Cys Phe Lys Leu Val Pro Val Glu Pro Glu Lys Val Glu Glu Ala Asn Glu Gly Glu Asn Asn Cys Ala Ala His 25 Pro Met Ser Gln His Gly Ile Glu Asp Pro Glu Lys Glu Val Leu Glu Trp Arg Phe Asp Ser Lys Leu Ala Phe His His Val Ala Arg Glu Leu His Pro Glu Tyr Tyr Lys Asp Cys Ser (SEQ ID NO:14).

An additional embodiment of the present invention relates to another DNA molecule encoding optimized HIV-1 Nef wherein the amino terminal myristylation site and dileucine motif have been deleted, as well as comprising a tPA leader peptide. This DNA molecule, opt tpanef (LLAA) comprises an open reading frame which encodes a Nef protein containing a tPA leader sequence fused to amino acid residue 6-216 of HIV-1 Nef (jfrl), wherein Leu-174 and Leu-175 are substituted with Ala-174 and Ala-175 (Ala-195 and Ala-196 in this tPA-based fusion protein). The nucleotide

sequence comprising the open reading frame of opt tpanef (LLAA) is disclosed herein as SEQ ID NO:15, as shown below:

CATGGATGCA ATGAAGAGAG GGCTCTGCTG TGTGCTGCTG CTGTGTGGAG CAGTCTTCGT
TTCGCCCAGC GAGATCTCCT CCAAGAGGTC CGTGCCCGGC TGGTCCACCG TGAGGGAGAG
GATGAGGAGG GCCGAGCCCG CCGCCGACAG GGTGAGGAGG ACCGAGCCCG CCGCCGTGGG
CGTGGGCGCC GTGTCCAGGG ACCTGGAGAA GCACGGCGCC ATCACCTCCT CCAACACCGC
CGCCACCAAC GCCGACTGCG CCTGGCTGGA GGCCCAGGAG GACGAGGAGG TGGGCTTCCC
CGTGAGGCCC CAGGTGCCCC TGAGGCCCAT GACCTACAAG GGCGCCGTGG ACCTGTCCCA
CTTCCTGAAG GAGAAGGGCG GCCTGGAGGG CCTGATCCAC TCCCAGAAGA GGCAGGACAT
CCTGGACCTG TGGGTGTACC ACACCCAGGG CTACTTCCCC GACTGGCAGA ACTACACCCC
CGGCCCCGGC ATCAGGTTCC CCCTGACCTT CGGCTGGTGC TTCAAGCTGG TGCCCGTGGA
GCCCGAGAAG GTGGAGGAGG CCAACGAGGG CGAGAACAAC TGCGCCGCCC ACCCCATGTC
CCAGCACGGC ATCGAGGACC CCGAGAAGGA GGTGCTGGAG TGGAGGTTCG ACTCCAAGCT
GGCCTTCCAC CACGTGGCCA GGGAGCTGCA CCCCGAGTAC TACAAGGACT GCTAAAGCCC
(SEQ ID NO:15).

5

10

15

The open reading frame of SEQ ID NO:7 encoding tPA-Nef (LLAA), disclosed herein as SEQ ID NO:16, is as follows:

Met Asp Ala Met Lys Arg Gly Leu Cys Cys Val Leu Leu Leu Cys Gly Ala Val Phe Val Ser Pro Ser Glu Ile Ser Ser Lys Arg Ser Val Pro 20 Gly Trp Ser Thr Val Arg Glu Arg Met Arg Arg Ala Glu Pro Ala Ala Asp Arg Val Arg Arg Thr Glu Pro Ala Ala Val Gly Val Gly Ala Val Ser Arg Asp Leu Glu Lys His Gly Ala Ile Thr Ser Ser Asn Thr Ala Ala Thr Asn Ala Asp Cys Ala Trp Leu Glu Ala Gln Glu Asp Glu Glu Val Gly Phe Pro Val Arg Pro Gln Val Pro Leu Arg Pro Met Thr Tyr Lys Gly Ala Val Asp Leu Ser His Phe Leu Lys Glu Lys Gly Gly Leu Glu Gly Leu Ile His Ser Gln Lys Arg Gln Asp Ile Leu Asp Leu Trp Val Tyr His Thr Gln Gly Tyr Phe Pro Asp Trp Gln Asn Tyr Thr Pro Gly Pro Gly Ile Arg Phe Pro Leu Thr Phe Gly Trp Cys Phe Lys Leu Val Pro Val Glu Pro Glu Lys Val Glu Glu Ala Asn Glu Gly Glu Asn 30 Asn Cys Ala Ala His Pro Met Ser Gln His Gly Ile Glu Asp Pro Glu Lys Glu Val Leu Glu Trp Arg Phe Asp Ser Lys Leu Ala Phe His His Val Ala Arg Glu Leu His Pro Glu Tyr Tyr Lys Asp Cys (SEQ ID NO:16). An adenoviral vector of the present invention may comprise a DNA sequence, regardless of codon usage, which expresses a wild type or modified Nef protein as 35 described herein, including but not limited to modified Nef proteins which comprise a deletion or substitution of Gly 2, a deletion of substitution of Leu 174 and Leu 175

and/or inclusion of a leader sequence. Therefore, partial or fully codon optimized DNA vaccine expression vector constructs are preferred since such constructs should result in increased host expression. However, it is within the scope of the present invention to utilize "non-codon optimized" versions of the constructs disclosed herein, especially modified versions of HIV Nef which are shown to promote a substantial cellular immune response subsequent to host administration.

Figure 20A-C show nucleotide sequences at junctions between nef coding sequence and plasmid backbone of nef expression vectors V1Jns/nef (Figure 20A), V1Jns/nef(G2A,LLAA) (Figure 20B), V1Jns/tpanef (Figure 20C) and V1Jns/tpanef(LLAA) (Figure 20C, also). 5' and 3' flanking sequences of codon optimized nef or codon optimized nef mutant genes are indicated by bold/italic letters; nef and nef mutant coding sequences are indicated by plain letters. Also indicated (as underlined) are the restriction endonuclease sites involved in construction of respective nef expression vectors. V1Jns/tpanef and V1Jns/tpanef(LLAA) have identical sequences at the junctions.

Figure 21 shows a schematic presentation of nef and nef derivatives. Amino acid residues involved in Nef derivatives are presented. Glycine 2 and Leucine 174 and 175 are the sites involved in myristylation and dileucine motif, respectively.

20 EXAMPLE 19

10

15

25

30

MRKAd5Pol Construction and Virus Rescue

Steps performed in the construction of the vectors, including the pre-adenovirus plasmid - Key steps performed in the construction of the vectors, including the pre-adenovirus plasmid denoted MRKAd5pol, is depicted in Figure 22. Briefly, the adenoviral shuttle vector for the full-length inactivated HIV-1 pol gene is as follows. The vector MRKpdelE1(Pac/pIX/pack450)+CMVmin+BGHpA(str.) is a derivative of the shuttle vector used in the construction of the MRKAd5gag adenoviral pre-plasmid. The vector contains an expression cassette with the hCMV promoter (no intronA) and the bovine growth hormone polyadenylation signal. The expression unit has been inserted into the shuttle vector such that insertion of the gene of choice at a unique BgIII site will ensure the direction of transcription of the transgene will be Ad5 E1 parallel when inserted into the MRKpAd5(E1-/E3+)Cla1 (or MRKpAdHVE3) pre-plasmid. The vector, similar to the original shuttle vector contains the Pac1 site, extension to the packaging signal region, and extension to the pIX gene. The synthetic full-length codon-optimized HIV-1 pol gene was isolated directly from the plasmid pV1Ins-HIV-pol-inact(opt). Digestion of this plasmid with BgI II releases the pol

gene intact (comprising a codon optimized IA pol sequence as disclosed in SEQ ID NO:3). The pol fragment was gel purified and ligated into the MRKpdelE1(Pac/pIX/pack450)+CMVmin+BGHpA(str.) shuttle vector at the BgIII site. The clones were checked for the correct orientation of the gene by using restriction enzymes DraIII/Not1. A positive clone was isolated and named MRKpdel+hCMVmin+FL-pol+bGHpA(s). The genetic structure of this plasmid was verified by PCR, restriction enzyme and DNA sequencing. The pre-adenovirus plasmid was constructed as follows. Shuttle plasmid MRKpdel+hCMVmin+FLpol+bGHpA(S) was digested with restriction enzymes Pac1 and Bst1107 I (or its isoschizomer, BstZ107 I) and then co-transformed into E. coli strain BJ5183 with linearized (Cla1 digested) adenoviral backbone plasmid, MRKpAd(E1-/E3+)Cla1. The resulting pre-plasmid originally named MRKpAd+hCMVmin+FLpol+bGHpA(S)E3+ is now referred to as "pMRKAd5pol". The genetic structure of the resulting pMRKAd5pol was verified by PCR, restriction enzyme and DNA sequence analysis. The vectors were transformed into competent E. coli XL-1 Blue for preparative production. The recovered plasmid was verified by restriction enzyme digestion and DNA sequence analysis, and by expression of the pol transgene in transient transfection cell culture. The complete nucleotide sequence of this pMRKAd5HIV-1pol adenoviral vector is shown in Figure 26 A-AO.

10

15

20

25

30

35

Generation of research-grade recombinant adenovirus - The pre-adenovirus plasmid, pMRKAd5pol, was rescued as infectious virions in PER.C6® adherent monolayer cell culture. To rescue infectious virus, 12  $\mu$ g of pMRKAd5pol was digested with restriction enzyme Pacl (New England Biolabs) and 3.3  $\mu$ g was transfected per 6 cm dish of PER.C6® cells using the calcium phosphate co-precipitation technique (Cell Phect Transfection Kit, Amersham Pharmacia Biotech Inc.). Pacl digestion releases the viral genome from plasmid sequences allowing viral replication to occur after entry into PER.C6® cells. Infected cells and media were harvested 6 -10 days post-transfection, after complete viral cytopathic effect (CPE) was observed. Infected cells and media were stored at  $\leq$  -60°C. This pol containing recombinant adenovirus is referred to herein as "MRKAd5pol". This recombinant adenovirus expresses an inactivated HIV-1 Pol protein as shown in SEQ ID NO:6.

#### **EXAMPLE 20**

#### MRKAd5Nef Construction and Virus Rescue

Construction of vector: shuttle plasmid and pre-adenovirus plasmid - Key steps performed in the construction of the vectors, including the pre-adenovirus

plasmid denoted MRKAd5nef, is depicted in Figure 23. Briefly, as shown in Example 19 above, the vector MRKpdelE1(Pac/pIX/pack450)+CMVmin+BGHpA(str.) is the shuttle vector used in the construction of the MRKAd5gag adenoviral pre-plasmid. It has been modified to contain the Pac1 site, extension to the packaging signal region, and extension to the 5 pIX gene. It contains an expression cassette with the hCMV promoter (no intronA) and the bovine growth hormone polyadenylation signal. The expression unit has been inserted into the shuttle vector such that insertion of the gene of choice at a unique Bgl11 site will ensure the direction of transcription of the transgene will be Ad5 E1 parallel when inserted into the MRKpAd5(E1-/E3+)Cla1 pre-plasmid. The synthetic 10 full-length codon-optimized HIV-1 nef gene was isolated directly from the plasmid pV1Jns/nef (G2A,LLAA). Digestion of this plasmid with Bgl11 releases the pol gene intact, which comprises the nucleotide sequence as disclosed in SEQ ID NO:13. The nef fragment was gel purified and ligated into the

MRKpdelE1+CMVmin+BGHpA(str.) shuttle vector at the Bgl11 site. The clones 15 were checked for correction orientation of the gene by using restriction enzyme Scal. A positive clone was isolated and named MRKpdelE1hCMVminFL-nefBGHpA(s). The genetic structure of this plasmid was verified by PCR, restriction enzyme and DNA sequencing. The pre-adenovirus plasmid was constructed as follows. Shuttle plasmid MRKpdelE1hCMVminFL-nefBGHpA(s) was digested with restriction 20 enzymes Pac1 and Bst1107 I (or its isoschizomer, BstZ107 I) and then co-transformed into E. coli strain BJ5183 with linearized (Cla1 digested) adenoviral backbone plasmid, MRKpAd(E1/E3+)Cla1. The resulting pre-plasmid originally named MRKpdelE1hCMVminFL-nefBGHpA(s) is now referred to as "pMRKAd5nef". The genetic structure of the resulting pMRKAd5nef was verified by PCR, restriction 25 enzyme and DNA sequence analysis. The vectors were transformed into competent E. coli XL-1 Blue for preparative production. The recovered plasmid was verified by restriction enzyme digestion and DNA sequence analysis, and by expression of the nef transgene in transient transfection cell culture. The complete nucleotide sequence of this pMRKAd5HIV-1nef adenoviral vector is shown in Figure 27A-AM. 30

Generation of research-grade recombinant adenovirus - The pre-adenovirus plasmid, pMRKAd5nef, was rescued as infectious virions in PER.C6<sup>®</sup> adherent monolayer cell culture. To rescue infectious virus, 12 μg of pMRKAdnef was digested with restriction enzyme Pac1 (New England Biolabs) and 3.3 μg was transfected per 6 cm dish of PER.C6<sup>®</sup> cells using the calcium phosphate coprecipitation technique (Cell Phect Transfection Kit, Amersham Pharmacia Biotech

35

Inc.). Pac1 digestion releases the viral genome from plasmid sequences allowing viral replication to occur after entry into PER.C6<sup>®</sup>cells. Infected cells and media were harvested 6-10 days post-transfection, after complete viral cytopathic effect (CPE) was observed. Infected cells and media were stored at ≤ -60°C. This nef containing recombinant adenovirus is now referred to as "MRKAd5nef".

#### **EXAMPLE 21**

5

10

15

20

25

30

35

# Construction of Murine CMV Promoter Containing Shuttle Vectors for Inactivated Pol and Nef/G2A,LLAA

The murine CMV (mCMV) was amplified from the plasmid pMH4 (supplied by Frank Graham, McMaster University) using the primer set: mCMV (Not I) Forward: 5'-ATA AGA ATG CGG CCG CCA TAT ACT GAG TCA TTA GG-3' (SEQ ID NO: 20); mCMV (Bgl II)Reverse: 5'-AAG GAA GAT CTA CCG ACG CTG GTC GCG CCT C-3' (SEQ ID NO:21). The underlined nucleotides represent the Not I and the Bgl II sites respectively for each primer. This PCR amplicon was used for the construction of the mCMV shuttle vector containing the transgene in the E1 parallel orientation. The hCMV promoter was removed from the original shuttle vector (containing the hCMV-gag-bGHpA transgene in the E1 parallel orientation) by digestion with Not I and Bgl II. The mCMV promoter (Not I/Bgl II digested PCR product) was inserted into the shuttle vector in a directional manner. The shuttle vector was then digested with Bgl II and the gag reporter gene (Bgl II fragment) was re-inserted back into the shuttle vector. Several clones were screened for correct orientation of the reporter gene. For the construction of the mCMV-gag in the E1 antiparallel orientation, the mCMV promoter was amplified from the plasmid pMH4 using the following primer set: mCMV (Asc I) Forward: 5'- ATA AGA ATG GCG CGC CAT ATA CTG AGT CAT TAG G (SEQ ID NO:22); mCMV (Bgl II) Reverse: 5' AAG GAA GAT CTA CCG ACG CTG GTC GCG CCT C (SEQ ID NO:23). The underlined nucleotides represent the Asc I and Bgl II sites, respectively for each primer. The shuttle vector containing the hCMV-gag transgene in the E1 antiparallel orientation was digested with Asc1 and Bgl11 to remove the hCMV-gag portion of the transgene. The mCMV promoter (Asc1/Bgl11 digested PCR product) was inserted into the shuttle vector in a directional manner. The vector was then digested with Bgl11 and the gag reporter gene (Bgl11 fragment) was re-inserted. Several clones were screened for correct orientation of the reporter gene. For each of the full length IA pol and full length nef/G2A,LLAA genes, cloning was performed using the unique

 $Bgl \ \Pi$  site within the mCMV-bGHpA shuttle vector. The pol and nef genes were excised from their respective pV1Jns plasmids by  $Bgl \ \Pi$  digestion.

#### **EXAMPLE 22**

Construction of mCMV Full Length Inactivated Pol and Full Length nef/G2A.LLAA Adenovectors

5

10

15

Each of these transgenes of Example 21 were inserted into the modified shuttle vector in both the E1 parallel and E1 anti-parallel orientations. Pac1 and BstZ110I digestion of each shuttle vector was performed and each specific transgene fragment containing the flanking Ad5 sequences was isolated and co-transformed with Cla I digested MRKpAd5(E3+) or MRKpAd5(E3-) adenovector plasmids via bacterial homologous recombination in BJ5183 E. coli cells. Recombinant preplasmid adenovectors containing the various transgenes in both the E3- and E3+ versions (and in the E1 parallel and E1 antiparallel orientations) were subsequently prepared in large scale following transformation into XL-1 Blue E. coli cells and analyzed by restriction analysis and sequencing.

#### **EXAMPLE 23**

Construction of hCMV-tpa-nef (LLAA) Adenovector

The tpa-nef gene was amplified out from GMP grade pV1Jns-tpanef (LLAA) vector using the primer sets: Tpanef (BamHI) F 5'-ATT GGA TCC ATG GAT GCA ATG AAG AGA GGG (SEQ ID 24); Tpanef (BamHI) R 5'-ATA GGA TCC TTA GCA GTC CTT GTA GTA CTC G (SEQ ID NO:25). The resulting PCR product was digested with BanHI, gel purified and cloned into the Bgl II site of MRKAd5CMV-bGHpA shuttle vector (Bgl II digested and calf intestinal phosphatase treated). Clones containing the tpanef (LLAA) gene (see SEQ ID NO:15 for complet coding region) in the correct orientation with respect to the hCMV promoter were selected following Sca I digestion. The resulting MRKAd5tpanef shuttle vector was digested with Pac I and Bst Z1101 and cloned into the E3+ MRKAd5 adenovector via bacterial homologous recombination techniques.

#### **EXAMPLE 24**

Immunogenicity of MRKAd5pol and MRKAd5nef Vaccine

Materials and Methods - Rodent Immunization - Groups of N=10 BALB/c

mice were immunized i.m. with the following vectors: (1) MRKAd5hCMV-IApol

(E3+) at either 10^7 vp and 10^9 vp; and (2) MRKAd5hCMV-IApol (E3-) at either

10^7 vp and 10^9 vp. At 7 weeks post dose, 5 of the 10 mice per cohort were boosted with the same vector and dose they initially received. At 3 weeks post the second does, sera and spleens were collected from all the animals for RT ELISA and IFNg ELIspot analyses, respectively. For all rodent immunizations, the Ad5 vectors were diluted in 5 mM Tris, 5% sucrose, 75 mM NaCl, 1 mM MgCl2, 0.005% polysorbate 80, pH 8.0. The total dose was injected to both quadricep muscles in 50 µL aliquots using a 0.3-mL insulin syringe with 28-1/2G needles (Becton-Dickinson, Franklin Lakes, NJ).

Groups of N=10 C57/BL6 mice were immunized i.m. with the following vectors: (1) MRKAd5hCMV-nef(G2A,LLAA) (E3+) at either 10^7 vp and 10^9 vp; (2) MRKAd5mCMV-nef(G2A,LLAA) (E3+) at either 10^7 vp and 10^9 vp; and (3) MRKAd5mCMV-tpanef(LLAA) (E3+) at either 10^7 vp and 10^9 vp. At 7 weeks post dose, 5 of the 10 mice per cohort were boosted with the same vector and dose they initially received. At 3 weeks post the second does, sera and spleens were collected from all the animals for RT ELISA and IFNg ELIspot analyses, respectively.

10

15

20

25

30

35

Non-human Primate immunization - Cohorts of 3 rhesus macaques (2-3 kg) were vaccinated with the following Ad vectors: (1) MRKAd5hCMV-IApol (E3+) at either 10^9 vp and 10^11 vp dose; and (2) MRKAd5hCMV-IApol (E3-) at either 10^9 vp and 10^11 vp; (3) MRKAd5hCMV-nef(G2A,LLAA) (E3+) at either 10^9 vp and 10^11 vp; and (4) MRKAd5mCMV-nef(G2A,LLAA) (E3+) at either 10^9 vp and 10^11 vp. The vaccine was administered to chemically restrained monkeys (10 mg/kg ketamine) by needle injection of two 0.5 mL aliquots of the Ad vectors (in 5 mM Tris, 5% sucrose, 75 mM NaCl, 1 mM MgCl<sub>2</sub>, 0.005% polysorbate 80, pH 8.0) into both deltoid muscles. The animals were immunized twice at a 4 week interval (T=0, 4 weeks).

Murine anti-RT and anti-nef ELISA - Anti-RT titers were obtained following standard secondary antibody-based ELISA. Maxisorp plates (NUNC, Rochester; NY) were coated by overnight incubation with 100 μL of 1 μg/mL HIV-1 RT protein (Advanced Biotechnologies, Columbia, MD) in PBS. For anti-nef ELISA, 100 uL of 1 ug/mL HIV-1 nef (Advanced Biotechnologies, Columbia, MD) was used to coat the plates. The plates were washed with PBS/0.05% Tween 20 using Titertek MAP instrument (Hunstville, AL) and incubated for 2 h with 200 μL/well of blocking solution (PBS/0.05% tween/1% BSA). An initial serum dilution of 100-fold was performed followed by 4-fold serial dilution. 100-μL aliquots of serially diluted samples were added per well and incubated for 2 h at room temperature. The plates

were washed and 100  $\mu$ L of 1/1000-diluted HRP-rabbit anti-mouse IgG (ZYMED, San Francisco, CA) were added with 1 h incubation. The plates were washed thoroughly and soaked with 100  $\mu$ L 1,2-phenylenediamine dihydrochloride/hydrogen peroxide (DAKO, Norway) solution for 15 min. The reaction was quenched by adding 100  $\mu$ L of 0.5M H<sub>2</sub>SO4 per well. OD<sub>492</sub> readings were recorded using Titertek Multiskan MCC/340 with S20 stacker. Endpoint titers were defined as the highest serum dilution that resulted in an absorbance value of greater than or equal to 0.1 OD<sub>492</sub> (2.5 times the background value).

Non-human primate and murine ELIspot assays - The enzyme-linked immuno-spot (ELISpot) assay was utilized to enumerate antigen-specific INFγ-secreting cells from mouse spleens (Miyahira, et al.1995, J. Immunol. Methods 181:45-54) or macaque PBMCs. Mouse spleens were pooled from 5 mice/cohort and single cell suspensions were prepared at 5x10<sup>6</sup>/mL in complete RPMI media (RPMI1640, 10% FBS, 2mM L-glutamine, 100U/mL Penicillin, 100 u/mL streptomycin, 10 mM Hepes, 50 uM β-ME). Rhesus PBMCs were prepared from 8-15 mL of heparinized blood following standard Ficoll gradient separation (Coligan, et al, 1998, Current Protocols in Immunology. John Wiley & Sons, Inc.). Multiscreen opaque plates (Millipore, France) were coated with 100 μL/well of either 5 μg/mL purified rat anti-mouse IFN-γ IgG1, clone R4-6A2 (Pharmingen, San Diego, CA), or 15 ug/mL mouse anti-human IFN-γ IgG2a (Cat. No. 1598-00, R&D Systems, Minneapolis, MN) in PBS at 4°C overnight for murine or monkey assays, respectively. The plates were washed with PBS/penicillin/streptomycin and blocked with 200 μL/well of complete RPMI media for 37 °C for at least 2 h.

To each well, 50 μL of cell samples (4-5x10<sup>5</sup> cells per well) and 50 μL of the antigen solution were added. To the control well, 50 μL of the media containing DMSO were added; for specific responses, either selected peptides or peptide pools (4 ug/mL per peptide final concentration) were added. For BALB/c mice immunized with the pol constructs, stimulation was conducted using a pool of CD4<sup>+</sup>-epitope containing 20-mer peptides (aa21-40, aa411-430, aa641-660, aa731-750, aa771-790) or a pool of CD8<sup>+</sup>-epitope containing peptides (aa201-220, aa311-330, aa781-800). For C57/BL6 mice immunized with the nef construct, either aa51-70 (CD8<sup>+</sup> T cell epitope) or aa81-100 (CD4<sup>+</sup>) peptide derived from the nef sequence was added for specific stimulation. In monkeys, the responses against pol were evaluated using two pools (L and R) of 20-aa peptides that encompass the entire pol sequence and overlap by 10 amino acids. In monkeys vaccinated with the nef constructs, a single pool containing 20-mer peptides covering the entire HIV-1 nef sequence and overlapping

by 10 aa was used. Each sample/antigen mixture was performed in triplicate wells for murine samples or in duplicate wells for rhesus PBMCs. Plates were incubated at 37°C, 5% CO<sub>2</sub>, 90% humidity for 20-24 h. The plates were washed with PBS/0.05% Tween 20 and incubated with 100 μL/well of either 1.25 μg/mL biotin-conjugated rat anti-mouse IFN-γ mAb, clone XMG1.2 (Pharmingen) or of 0.1 ug/mL biotinylated anti-human IFN-gamma goat polyclonal antibody (R&D Systems) at 4°C overnight. The plates were washed and incubated with 100 μL/well 1/2500 dilution of strepavidin-alkaline phosphatase conjugate (Pharmingen) in PBS/0.005% Tween/5% FBS for 30 min at 37 °C. Spots were developed by incubating with 100 μL/well 1-step NBT/BCIP (Pierce Chemicals) for 6-10 min. The plates were washed with water and allowed to air dry. The number of spots in each well was determined using a dissecting microscope and the data normalized to 10<sup>6</sup> cell input.

Non-human Primate anti-RT ELISA - The pol-specific antibodies in the monkeys were measured in a competitive RT EIA assay, wherein sample activity is determined by the ability to block RT antigen from binding to coating antibody on the plate well. Briefly, Maxisorp plates were coated with saturating amounts of pol positive human serum (#97111234). 250 uL of each sample is incubated with 15 uL of 266 ng/mL RT recombinant protein (in RCM 563, 1% BSA, 0.1% tween, 0.1% NaN<sub>3</sub>) and 20 uL of lysis buffer (Coulter p24 antigen assay kit) for 15 min at room temperature. Similar mixtures are prepared using serially diluted samples of a standard and a negative control which defines maximum RT binding. 200 uL/well of each sample and standard were added to the washed plate and the plate incubated 16-24 h at room temperature. Bound RT is quantified following the procedures described in Coulter p24 assay kit and reported in milliMerck units per mL arbitrarily defined by the chosen standard.

Results - Rodent Studies - BALB/c mice (n=5 mice/cohort) were immunized once or twice with varying doses of MRKAd5hCMV-IApol(E3+) and MRKAd5hCMV-IApol(E3-). At 3 weeks after the second dose, Anti-pol IgG levels were determined by an ELISA assay using RT as a surrogate antigen. Cellular response were quantified via IFNy ELISpot assay against pools of pol-epitope containing peptides. The results of these assays are summarized in Table 10. The results indicate that the mouse vaccinees exhibited detectable anti-RT IgGs with an adenovector dose as low as 10^7 vp. The humoral responses are highly dose-dependent and are boostable with a second immunization. One or two doses of either pol vectors elicit high frequencies of antigen-specific CD4<sup>+</sup> and CD8<sup>+</sup> T cells; the responses are weakly dose-dependent but are boostable with a second immunization.

Table 10. Immunogenicity of MRKAd5pol Vectors in BALB/c mice.

|       |                       |         |                 | An   | ti-RT IgG Tite  | rs*            | S             | FC/10^6 cell            | s°                      |
|-------|-----------------------|---------|-----------------|--|-----------------|----------------|---------------|-------------------------|-------------------------|
| Group | Vaccine               | Dese    | No. of<br>Doses | GMT  | +SE             | -SE            | Medlum        | CD4+<br>peptide<br>pool | CD8+<br>peptide<br>pool |
| 1     | MRKAdShCMVFLpal (E3+) | 10^7 vp | 2<br>1          | 310419<br>919                                | 301785<br>372   | 153020<br>265  | 1(1)<br>1(1)  | 75(4)<br>72(9)          | 2313(67)<br>533(41)     |
| 2     | MRKAdShCMVFLpol (E3+) | 10^9 vp | 2               | 1638400 <sup>b</sup><br>713155               | 0<br>528520     | 0<br>303555    | 2(2)<br>1(1)  | 114(9)<br>48(7)         | 2083(182)<br>733(89)    |
| 3     | MRKAd5hCMVFLpol (E3-) | 10^7 vp | 2               | 310419<br>6400                               | 388218<br>14013 | 172097<br>4393 | O(0)<br>10(8) | 223(7)<br>141(21)       | 2607(27)<br>409(28)     |
| 4     | MRKAd5hCMVFLpol (E3-) | 10^9 vp | 2               | 1638400 <sup>b</sup><br>1241675 <sup>b</sup> | 0<br>396725     | 0<br>300681    | 1(1)<br>0(0)  | 160(13)<br>39(13)       | 2385(11)<br>833(83)     |
|       | Naive                 | none    | none            | 57   | 9               | 7              | 9(2)          | 11(4)                   | 10(1)                   |

GMT, geometric mean titer of the cohort of 5 mice; SE, standard error of the gemetric mean

5 C57/BL6 mice were immunized once or twice with varying doses of MRKAd5hCMV-nef(G2A,LLAA) (E3+), MRKAd5mCMV-nef(G2A,LLAA) (E3+) at either 10^7 vp and(3) MRKAd5mCMV-tpanef(LLAA) (E3+) at either 10^7 vp and 10^9 vp. The immune response were analyzed using similar protocols and the results are listed in Table 11. While anti-nef IgG responses could not be detected in this model system with any of the constructs, there are strong indications of a cellular immunity generated against nef using the ELIspot assay.

Table 11. Immunogenicity of MRKAd5nef Vectors in C57/BL6 mice.

|       |                       |         |                 | Ап  | ti-nef IgG Tite | ers" | S      | FC/10^6 cell    | <b>3</b> b       |
|-------|-----------------------|---------|-----------------|-----|-----------------|------|--------|-----------------|------------------|
| Group | Vaccine               | Dose    | No. of<br>Doses | GMT | +SE             | -SE  | Medium | 2251-70<br>CD8+ | 8881-100<br>CD4+ |
|       | MRKAd5hCMVFLnel (E3+) | 10^7 vp | 2               | 174 | 70              | 50   | 1(1)   | 23(1)           | 1(1)             |
| •     |                       |         | 1               | 132 | 42              | 32   | 0(0)   | 0(0)            | 0(0)             |
|       | MRKAd5hCMVFLnel (E3+) | 10^9 Vp | 2               | 174 | 70              | 50   | 0(0)   | 61(7)           | 4(2)             |
| •     | in 100                |         | 1               | 132 | 42              | 32   | 1(1)   | 62(7)           | 3(1)             |
| 3     | MRKAd5mCMVFLnet (E3+) | 10^7 vp | 2               | 132 | 42              | 32   | 3(1)   | 15(5)           | 5(2)             |
|       | , , ,                 |         | 1 -             | 115 | 46              | 33   | 3(2)   | 3(2)            | 4(2)             |
| · ·   | MRKAd5mCMVFLnel (E3+) | 10^9 vp | 2               | 132 | 42              | 32   | 4(2)   | 83(13)          | 5(1)             |
| •     |                       |         | 1               | 132 | 42              | 32   | 2(1)   | 28(2)           | 4(0)             |
| 5     | MRKAd5mCMVtpanel(E3+) | 10^7 VD | 2               | 132 | 42              | 32   | 3(2)   | 14(2)           | 5(1)             |
| •     |                       | 1       | 1               | 100 | 0               | 0    | 3(1)   | 13(4)           | 10(3)            |
| 6     | MRKAd5mCMVtpanet(E3+) | 10/9 VD | 2               | 230 | 170             | 98   | 3(2)   | 145(29)         | 4(0)             |
|       |                       |         | 1               | 115 | 46              | 33   | 7(1)   | 151(14)         | 10(0)            |
|       | Naive                 | none    | none            | 152 | 78              | 52 · | 21(2)  | - 18(6)         | 28(3)            |

<sup>&</sup>quot;GMT, geometric mean titler of the cohort of 5 mice; SE, standard error of the gemetric mean

15

Monkey Studies - Cohorts of 3 rhesus macaques were immunized with 2 doses of MRKAd5hCMV-IApol(E3+) and MRKAd5hCMV-IApol(E3-). The number of antigen-specific T cells (per million PBMCs) were enumerated using one of two

Near or at the upper limit of the serial dilution; hence, could be greater than this value

<sup>°</sup>No. of Spot-forming Cells per millon splecnoytes; mean values of triplicates are reported along with standard errors in parenthesis.

No. of spol-forming calls per million spiechoytes; mean values of triplicates are reported along with standard errors in parenthesis.

peptide pools (L and R) that cover the entire pol sequence; the results are listed in Table 12. Moderate-to-strong T cell responses were detected in the vaccinees using either constructs even at a low dose of 10^9 vp. Longitudinal analyses of the anti-RT antibody titers in the animals suggest that the pol transgene product is expressed efficiently to elicit a humoral response (Table 13). It would appear that generally higher immune responses were observed in animals that received the E3- construct compared to the E3+ virus.

Table 12. Pol-specific T Cell Responses in MRKAd5pol Immunized Rhesus

10 Macaques.

5

| Vaccine (T=0,4 wks)    | Monk #       |      | Prebleec |      |      | T=4   |       |      | T=7   |       |      | T=16  |       |
|------------------------|--------------|------|----------|------|------|-------|-------|------|-------|-------|------|-------|-------|
|                        |              | Mock | Pol L    | PolR | Mock | Pol L | Pol R | Mock | Pol L | Pol R | Mock | Pol L | Pol R |
| MRKAd5hCMV-IApod(E3+)  | 990100       | 1    | 0        | 0    | 1    | 38    | 31    | 0    | 52    | 146   | 0    | 49    | 715   |
| 10^11 VD               | 99C215       | 1 1  | 2        | 2    | 10   | 98    | 249   | 1    | 109   | 305   | 22   | 88    | 250   |
|                        | 99D201       | 5    | 5        | 4    | 6    | 149   | 85    | 0    | 40    | 35    | ٥    | 35    | 18    |
| MRKACEHOMV-IApol(E3+)  | 99D212       | 0    | 2        | 0    | 4    | 331   | 114   | 0    | 58    | 14    | 0    | 6     | 6     |
| 10/9 Vp                | 99D18D       | 0    | 4        | 2    | 0    | 19    | 192   | 4    | 38    | 158   | 5    | 38    | 108   |
|                        | 99C201       | 8    | 5        | 21   | 8    | 62    | 62    | ٥    | 18    | 32    | ١,   | 14    | 65    |
| MRKACCHOMV-I Apod(E3-) | 99D239       | 5    | 2        | 2    | 20   | 82    | 172   | 1    | 68    | 114   | 9    | 21    | - 40  |
| 1041 vp                | 99C186       | 4    | 12       | 6    | 5    | 120   | 421   | 2    | 271   | 489   | 16   | 875   | 530   |
|                        | 99C084       | 1    | 8        | 9    | 8    | 84    | 464   | ٥    | 14    | 238   | 1    | 24    | 284   |
| MRKAd5hCMV-IApol(E3-)  | 007C         | 10   | 10       | 8    | 12   | 724   | 745   | 4    | 322   | 376   | 4    | 188   | 176   |
| 10/9 vp                | CDIG         | 2    | 0        | 1    | 5    | 474   | 468   | 0    | 232   | 212   | 0    | 101   | 121   |
| •                      | <b>CD</b> 11 | 6    | 6        | 12   | 10   | 98    | 110   | 5    | 60    | 80    | 8    | 25    | 34    |
| Nave                   | 083Q         | nd   | nd       | nd   | nd   | nd    | nd    | 4    | 2     | 2_    | 2    | 1     | 2     |

nd, not determined Renorted are SFC per million PBMCs; mean of duplicate wells.

Table 13. Anti-RT Ig Levels in MRKAd5pol Immunized macaques.

| RT ANTIBODY ASSAYTITERS IN MMU  | mL  |      |       |      |
|---------------------------------|-----|------|-------|------|
| Vaccine/Monkey Tag              | T=4 | T=7  | T=12  | T=16 |
| MRKAd5hCMV-IApol(E3+), 10^11 vp |     |      |       |      |
| 99C100                          | 61  | 1999 | 5928  | 4768 |
| 99C215                          | 81  | 1541 | 2356  | 2767 |
| 99D201                          | 53  | 336  | 539   | 387  |
| MRKAd5hCMV-IApol(E3+), 10^9 vp  |     |      |       |      |
| 99D212                          | 10  | 40   | 49    | 68   |
| 99D180                          | <10 | 36   | 79    | 93   |
| 99C201                          | <10 | 37   | 71    | 76   |
| MRKAd5hCMV-IApol(E3-), 10^11 vp |     |      | ·     |      |
| 99D239                          | 44  | 460  | 1234  | 1015 |
| 99C186                          | 21  | 233  | 480   | 345  |
| 990084                          | 235 | 2637 | 2858_ | 1626 |
| MRKAd5hCMV-lApol(E3-), 10^9 vp  |     |      |       |      |
| CC7C                            | 32  | 175  | 306   | 235  |
| @16                             | 20  | 140  | 273   | 419  |
| <b>©</b> 11                     | 15  | 112_ | 149   | 237  |

When rhesus macaques were immunized i.m. with two doses of MRKAd5nef constructs, vigorous T cell responses ranging from 100 to as high as 1100 per million were observed in 8 of 12 vaccinees (Table 14). The efficacies of the mCMV- and hCMV- driven nef constructs are comparable on the basis of the data generated thus far.

Table 14. Nef-specific T cell Responses in MRKAd5nef Immunized Rhesus

Macaques.

10

15

20

25

| Vaccine (T=0,4 wks)            | Monk # | P    | ne  | Ţ:   | <b>=4</b> | Ţ.   | <u>:</u> 7 | T□   | 16   |
|--------------------------------|--------|------|-----|------|-----------|------|------------|------|------|
|                                |        | Mock | Nef | Mock | Nef       | Mock | Net        | Mock | Nef  |
| MRKAd5hCMV-nef(G2A,LLAA) (E3+) | CD2D   | 0    | 4   | 31   | 440       | 4    | 368        | 1    | 251  |
| 10^11 vp                       | CC7B   | 0    | 0   | 2    | 521       | 0    | 178        | 1    | 1522 |
|                                | CC61   | 2    | 9   | 31   | 112       | 0    | 108        | .11  | 100  |
| MRKAd5hCMV-nef(G2A,LLAA) (E3+) | CC2K   | 9    | 9   | 6    | 52        | 0    | 35         | 0    | 15   |
| 10^9 VD                        | CD15   | 5    | 4   | 30   | 998       | 2    | 586        | 0    | 434  |
| ·                              | CD16   | 6    | 1   | 6    | 1146      | 0    | 369        | 1    | 212  |
| MRKAd5mCMV-nef(G2A,LLAA) (E3+) | 99D191 | 1    | 5   | 4    | 614       | 0    | 298        | 2    | 419  |
| 10^11 vp                       | 99D144 | 4    | 6   | 5    | 434       | 0    | 1100       | 2    | 932  |
| •                              | 99C193 | 1 1  | 2   | 1    | 58        | 1    | 22         | 0    | 64   |
| MRKAd5mCMV-net(G2A,LLAA) (E3+) | 99D224 | 1    | 11  | 14   | 231       | 1    | 125        | 0    | 70   |
| 10^9 vp                        | 99D250 | 8    | . 9 | 4    | 108       | 0    | 54         | 0    | 5    |
| •                              | 990120 | 1    | 6   | 20   | 299       | 0    | 92         | 0    | 79   |
| Naive                          | 083Q   | nd   | nd  | 1B   | 22        | 4    | 5          | 2    | 1    |

## **EXAMPLE 25**

Comparison of Clade B vs. Clade C T Cell Responses in HIV-Infected Subjects

PBMC samples collected from two dozens of patients infected with HIV-1 in

US were tested in ELISPOT assays with peptide pools of 20-mer peptides overlapping
by 10 amino acids. Four different peptide pools were tested for cross-clade
recognition, and they were either derived from a clade B-based isolate (gag H-b; nefb) or a clade C-based isolate (gag H-c, nef-c). Data in Table 15 shows that T cells
from these patients presumably infected with clade B HIV-1 could recognize clade C
gag and nef antigens in ELISPOT assay. Correlation analysis further demonstrated
that these T cell responses against clade C gag peptide pool were about 60% of the
clade B counterpart (Figure 24), while the T cell responses against clade C nef were
about 85% of the clade B counterpart (Figure 25). These results suggest that cellular
immune responses generated in patients infected with clade B HIV-1 can recognize
gag and nef antigens derived from clade C HIV-1. These data show that a HIV
vaccine, such as a DNA or MRKAd5-based adenoviral vaccine expressing a clade B

gag and/or nef antigen will potentially have the ability to provide a prophylactic and/or therapetic advantage on a global scale.

5

Table 15
Responses Shown as the Number of gIFN-Secreting T Cells per Million PBMCs

| subject | bleed date | gag epitope # | mock | gag H-b | gagH-c | nef-b | nef-c |
|---------|------------|---------------|------|---------|--------|-------|-------|
|         |            | from mapping) |      |         |        |       |       |
| #100    | 19-Jul-99  | 12            | 10   | 3950    | 1385   | 1295  | 1300  |
| #101    | 25-Jul-99  | 3             | 15   | 3885    | 1280   | na    | 1020  |
| #102    | 25-Jul-99  | 4             | 15   | 1740    | 850    | 1255  | 1785  |
| #104    | 7-Jun-99   | 2             | . 5  | 1355    | 1185   | na    | 1060  |
| #107    | 11-Oct-99  | 2             | 25   | 3305    | 2795   | 670   | 870   |
| #405    | 11-Jul-99  | 2             | 15   | 4575    | 3180   | 1700  | 1500  |
| #501    | 19-Jul-99  | 2             | 15   | 1100    | 570    | 3365  | 3460  |
| #505    | 18-Jul-99  | 5             | 10   | 2145    | 1725   | 1235  | na    |
| #506    | 28-Feb-99  | 2             | 25   | 150     | 45     | 400   | 610   |
| #701    | 28-Mar-99  | 5             | 30   | 7620    | 4775   | 3320  | 2780  |
| #709    | 17-May-99  | 3             | 15   | 2785    | 1945   | 1090  | 1630  |
| #710    | 24-May-99  | 4             | 5    | 1055    | 1080   | 2210  | 2140  |
|         | 1          | -             |      |         |        |       |       |

10

15

EXAMPLE 26
Characterization and Production of MRKAd5pol and MRKAd5nef

Vectors in Roller Bottles

Expansion of nef and pol Adenovectors - Nef and pol CsCl purified MRKAd5 seeds were used to infect roller bottles to produce P4 virus to be used as a seed for further experiments. P4 MRKAd5 pol and nef vectors were used to infect roller bottles at an MOI 280 vp/cell, except for hCMV-tpa-nef [E3+] which was infected at an MOI of 125 due to low titers of seed obtained at P4.

Table 16 Viral particle concentrations for P5 nef and pol adenovectors

| Adenovector        | AEX Titer (10 <sup>10</sup> vp/ml culture) | AEX Titer<br>(10 <sup>4</sup> vp/cell) | Amplification<br>Ratio |
|--------------------|--|--|------------------------|
| hCMV-FL-nef [E3+]  | 1.1  | 0.9                                    | 30                     |
| mCMV-FL-nef [E3+]  | 2.2  | 2.1                                    | 75                     |
| hCMV-tpa-nef [E3+] | 0.07                                       | 0.1                                    | 5                      |
| mCMV-tpa-nef [E3+] | 1.3  | 0.9                                    | 35                     |
| hCMV-FL-pol [E3+]  | 2.7  | 2.1                                    | 75                     |
| hCMV-FL-pol [E3-]  | 1.9  | 1.3                                    | 45                     |

5 Roller Bottle Passaging - Passaging of the pol and nef constructs continued through passage seven. Cell-associated (freeze/thaw lysis) and whole broth (tritonlysis) titers obtained in all passages were very consistent. In general, MRKAd5pol is ca. 70% as productive as MRKAd5gag while MRKAd5nef is ca. 25% as productive as MRKAd5gag. Samples of P7 virus for both constructs were analyzed by V&CB by restriction digest analysis and did not show any rearrangements.

Table 17. Passage Six Viral Productivity for MRKAd5pol and MRKAd5nef

| 14010             |            | Xviable (10<br>Viabile |           | Cell Passage | (Cell Associated)  | Titer          | Amplification | Triton Lysis Titer             |
|-------------------|------------|------------------------|-----------|--------------|--------------------|----------------|---------------|--------------------------------|
|                   |            | Infection              | Harvest   | Number       | 1010 vp/ml culture | 104 vp/cell    | Ratio         | 10 <sup>10</sup> vp/ml culture |
| hCMV-FL-nef [B3+] | pool       | 1.22, 85%              |           | 62           | 0.8                | 0.7            | 25            | 1.6                            |
|                   | <u>i</u>   |                        | 0.99, 62% |              |                    | <del>-</del> - |               |                                |
|                   | 2          |                        | 1.10, 72% | 1            | }                  |                |               |                                |
| bCMV-FL-pol [E3+] | pool       | 1.42, 89%              |           | 62           | 4.5                | 3.2            | 115           | 7.0                            |
|                   | <u>. i</u> |                        | 1.22, 70% |              |                    |                |               |                                |
|                   | 2          | [                      | 1.42, 74% |              |                    |                |               |                                |

15 Table 18. Passage Seven Viral Productivity for MRKAd5pol and MRKAd5nef

| Xviable / If College | ABX Titer | Titer | Amplification |

|                   |      | Xviable (10<br>Viabili<br>Infection | ) <sup>6</sup> cells/ml), | Cell Passage<br>Number | AEX Titer<br>(Cell Associated)<br>10 <sup>10</sup> vp/ml culture | Titer  10 <sup>4</sup> vp/cell | Amplification<br>Ratio | Triton Lysis Titer  10 <sup>m</sup> vp/ml culture |
|-------------------|------|-------------------------------------|---------------------------|------------------------|--|--------------------------------|------------------------|---|
| hCMV-FL-nef [E3+] | Pool | 1.33, 90%                           |                           | 66                     | 1.0  | 0.8                            | 29                     | 2.1   |
| 1                 | 1    |                                     | 0.96, 70%                 |                        |  | _                              |                        |   |
|                   | 2    |                                     | 1.18, 73%                 | .]                     | 1  |                                |                        |   |
| bCMV-FL-pol [E3+] | Pool | 0.90*, 90%                          |                           | 56                     | 4.2  | 4.7                            | 168                    | 6.5   |
|                   | 1    |                                     | 1.18, 88%                 |                        |  |                                |                        |   |
|                   | 2    |                                     | 1.04, 80%                 |                        |  |                                | <u> </u>               |   |

MRKAd5nef and MRKAd5pol Viral Production Kinetics - A timecourse experiment was carried out in roller bottles to determine if the viral production kinetics of the MRKAd5pol and MRKAd5nef vectors were similar to those of MRKAd5gag. PER.C6® cells in roller bottle cultures were infected at an MOI of 280 vp/cells with P5 MRKAd5pol, P5 MRKAd5nef and P7 MRKAd5gag; for each adenovector, two infected bottles were sampled at 24, 36, 48, and 60 hours post infection. In addition, two bottles were left unsampled until 48 hpi when they were harvested under the Phase I process conditions. The anion-exchange HPLC viral particle concentrations of the freeze-thaw recovered cell associated virus at the 24, 36,

48, and 60 hpi timepoints are shown in Figure 29A-B. The QPA titers show a similar trend (data not shown).

Comparison of hCMV- and mCMV-FL-nef - As the titers obtained with the MRKAd5nef construct (hCMV-FL-nef) were lower than those obtained with MRKAd5gag or MRKAd5pol, a viral productivity comparison experiment was performed with mCMV-FL-nef. For each of the two adenovectors (hCMV- and mCMV-FL-nef), two roller bottles were infected at an MOI of 280 vp/cell with passage five clarified lysate. The macroscopic and microscopic observations of the four roller bottles were identical at the time of harvest. Analysis of the clarified lysate produced indicated a higher viral particle concentration in the bottles infected with mCMV-FL-nef, as shown in Table 19. It is stipulated that the higher productivity with mCMV promoter driven nef vector is due to lower nef expression levels in PER.C6® cells- experiments are underway at V&CB to measure nef expression levels.

15
Table 19. Passage Six Viral Productivity Comparison of hCMV- and mCMV-FL-nef

|             | (    | Xv (10 <sup>6</sup> cells/m | l), Viability (%) | Cell Passage | AEX Titer                      | Titer       | Amplification | Triton Lysis Titer             |
|-------------|------|-----------------------------|-------------------|--------------|--------------------------------|-------------|---------------|--------------------------------|
|             |      | Infection                   | Harvest           | Number       | 10 <sup>10</sup> vp/ml culture | 104 vp/cell | Ratio         | 10 <sup>to</sup> vp/ml culture |
| hCMV-FL-nef | Pool | 1.11, 91%                   |                   | 60           | 1.5                            | 1.4         | 50            | 2.8                            |
| (MRKAd5nef) | . 1  |                             | 1.23,75%          |              |                                |             |               |                                |
| ľ           | 2    | •                           | 1.34, 74%         |              | [                              |             |               |                                |
| mCMV-FL-nef | Pool | 1.11, 91%                   |                   | 60           | 2.3                            | 2.1         | 75            | 4.6                            |
| 1           | - 1  |                             | 1.49, 84%         |              |                                |             |               |                                |
|             | 2    |                             | 1.18, 77%         |              |                                |             | ·             |                                |

20

25

30

#### **EXAMPLE 27**

# Characterization and Large Scale Production of MRKAd5nef Virus in Bioreactors

Materials and Methods - The experiment of the present example was run twice under the following conditions: 36.5°C, DO 30%, pH 7.30, 150rpm agitation rate, no sparging, Life Technologies (Gibco, Invitrogen) 293 SFM II (with 6mM L-glutamine), 0.5M NaOH as base for pH control. During the first run (B20010115), two 10L stirred vessel bioreactors were inoculated with PER.C6® cells at a concentration of 0.2x106 cells/ml. Cells were grown until they reached a cell concentration of approximately 1x106 cells/ml. The cells were infected with uncloned MRKAd5nef (G2A,LLAA) at a MOI of 280 virus particles (vp)/cell. For the second batch (B20010202), the same procedure as the first run was used, except the cells

were infected with cloned MRAd5nef. During both runs, the bioreactors were harvested 48 hours post-infection. Samples were taken and virus concentrations were determined from whole broth (with triton lysis), supernatant, and cell pellets (3 X freeze/thaw) with the AEX and QPA assays. Metabolites were measured with BioProfile 250 throughout the process.

Table 20: Experimental Conditions

| Temperature | - 36.5 ℃ |
|-------------|----------|
| DO          | 30%      |
| PH          | 7.30     |
| Agitation   | 150 rpm  |
| Sparging    | None     |

Table 21: Virus source used for experiments.

10

| Run | Batch ID    | Cloned/Uncloned<br>MRKAd5nef | MOI<br>(vp/cells) |
|-----|-------------|------------------------------|-------------------|
| #1  | B20010115-1 | Uncloned                     | 280               |
|     | B20010115-2 | Uncloned                     | 280               |
| #2  | B20010202-1 | Cloned                       | 280               |
|     | B20010202-2 | Cloned                       | 280               |

Results - Table 22 and 23 show an the ability to scale up production of MRKAd5nef by growth in a bioreactor.

15

Table 22: Virus Concentration as measured by the AEX assay

| Run | Batch ID    | Cloned/Uncloned | Virus Concentration @ 48hpi (1x10 <sup>13</sup> vp/L) |                  |       |               |  |  |
|-----|-------------|-----------------|---|------------------|-------|---------------|--|--|
|     |             | MRKAd5nef       | Supernatant   | Clarified Lysate | Total | Triton Lysate |  |  |
| #1  | B20010115-1 | Uncloned        | 0.72  | 3.26             | 3.98  | 5.76          |  |  |
| "-  | B20010115-2 | Uncloned        | 0.38  | 1.67             | 2.05  | 2.46          |  |  |
| #2  | B20010202-1 | Cloned          | 0.80  | 6.00             | 6.80  | 8.88          |  |  |
| "-  | B20010202-2 | Cloned          | 0.50  | 6.00             | 6.50  | 8.47          |  |  |

Table 23: Virus Titers as measured by the QPA assay

| Run | Batch ID    | Cloned/Uncloned | Virus Concentration @ 48hpi (1x10 <sup>11</sup> IU/L) |             |                     |       |                  |  |
|-----|-------------|-----------------|---|-------------|---------------------|-------|------------------|--|
|     | ·           | MRKAd5nef       | Whole<br>Broth  | Supernatant | Clarified<br>Lysate | Total | Triton<br>Lysate |  |
| #1  | B20010115-1 | Uncloned        | 0.13  | 1.12        | 1.76                | 2.88  | 11.28            |  |
| ,   | B20010115-2 | Uncloned        | 0.14  | 0.73        | 1.54                | 2.27  | 5.86             |  |
| #2  | B20010202-1 | Cloned          | 0.14  | 0.97        | 1.62                | 2.69  | 11.89            |  |
|     | B20010202-2 | Cloned          | 0.14  | 1.17        | 1.70                | 2.97  | 12.47            |  |

20

The present invention is not to be limited in scope by the specific embodiments described herein. Indeed, various modifications of the invention in addition to those described herein will become apparent to those skilled in the art

from the foregoing description. Such modifications are intended to fall within the scope of the appended claims.

### **EXAMPLE 28**

## MRKAd5HIV-1gag Boosting of DNA-Primed Animals

5

10

15

20

25

30

Groups of 3-5 rhesus macaques were immunized with (a) 5 mgs of V1Jns-Flgag (pVIJnsCMV(no intron)-FL-gag-bGHpA), (b) 5 mgs of V1Jns-Flgag formulated with 45 mgs of a non-ionic block copolymer CRL1005, or (c) 5 mgs of V1Jns-Flgag formulated with 7.5 mgs of CRL1005 and 0.6 mM benzalkonium chloride at weeks 0, 4, and 8. All animals received a single dose of 10e7 viral particles (vp) of the MRKAd5HIV-1gag at week 26. Note: 10e7 is too low to prime or boost effectively when used as a single modality (dose is selected to mimic preexposure to adenovirus); see Figure 32.

Blood samples were collected from all animals at several time points and peripheral blood mononuclear cells (PBMCs) were prepared using standard Ficoll method. The PBMCs were counted and analyzed for gamma-interferon secretion using the ELISpot assay (Table 24). For each monkey, the PBMCs were incubated overnight either in the absence (medium) or presence of a pool (called "gag H") of 50 20-aa long peptides that encompass the entire HIV-1 gag sequence.

The results indicate that MRKAd5HIV-1gag was very effective in boosting the T cell immune responses in these monkeys. At week 28 or 2 weeks after the viral boost, the number of gag-specific T cells per million PBMCs increased 2-48 fold compared to the levels observed at week 24 or 2 weeks prior to the boost.

The PBMCs were also analyzed by intracellular gamma-interferon staining prior to (at week 10) and after the MRKAd5gag boost (at week 30). The results for select animals are shown on Figure 31. The results indicate that (a) immunization with DNA/adjuvant formulation elicited T cell responses which can either be balanced, CD4<sup>+</sup>-biased or CD8<sup>+</sup>-biased, and (b) boosting with the MRKAd5gag construct produced in all cases a strongly CD8<sup>+</sup>-biased response. These results suggest that boosting with MRKAd5HIV-1gag construct is able to improve the levels of antigen-specific CD8<sup>+</sup> T cells.

25. 25. 25. 25. 25. 25. 25.

959 1915 1549 1549 2 8 2 8 2 8 E

38888 85888 និងខ 衰衰器至器 22 2 2 8 8 2 2 2 8 22 8 8 158 × 28 £ 284 135 198 36 36 8 <u>5</u> 0 8 5 名节路 85528 ㅁㅁ똧걸 Table 24, Boosting of DNA/Adjuvent-Primed Rhosus Monkeys with MRKAd5gag

Number of SFCmillion PBMCs

Greet Browst Names I T. CC1K CC1K CB6F AKBB COSK COSK AW39 AW20 CAAR CB68 CB6W CB70 CAL10057.5 mgs + 0.6 mM BAX

#### **EXAMPLE 29**

# Construction of gagpol fusion for MRKAd5gagpol fusion constructs

The open reading frames for the codon-optimized HIV-1 gag gene was fused directly to the open reading frame of the IA pol gene (consisting of RT, RNAseH and integrase domains) by stepwise PCR. Because the gene (SEQ ID NO: 38) does not include the protease gene and the frameshift sequence, it encodes a single polypeptide of the combined size of p55, RT, RNAse H and integrase (1350 amino acids; SEQ ID NO: 39).

5

10

15

20

25

30

35

The fragment that extends from the BstEII site within the gag gene to the last non-stop codon was ligated via PCR to a fragment that extends from the start codon of the IApol to a unique BamHI site. This fragment was digested with BstEII and BamHI. Construction of gag-IApol fusion was achieved via three-fragment ligation involving the PstI-BstEII gag digestion fragment, the BstEII/BamHI digested PCR product and long PstI/BamHI V1R-FLpol backbone fragment.

The MRKAd5-gagpol adenovirus vector was constructed using the BglII fragment of the V1R-gagpol containing the entire ORF of gag-IApol fusion gene.

#### **EXAMPLE 30**

Immunogenicity Studies in Non-Human Primates

Cohorts of three (3) macaques were immunized with 10e8 or 10e10 viral particles (vp) of one of the following MRKAd5 HIV-1 vaccines: (1) MRKAd5gag; (2) MRKAd5pol; (3) MRKAd5nef; (4) a mixture containing equal amounts of MRKAd5gag, MRKAd5pol, and MRKAd5nef, or (5) a mixture of equal amounts of MRKAd5gagpol and MRKAd5nef. The vaccines were administered at weeks 0 and

MRKAd5gagpol and MRKAd5nef. The vaccines were administered at weeks 0 and 4.

The T cell responses against each of the HIV-1 antigens were assayed by IFN-gamma ELISpot assay using pools of 20-aa peptides that encompass the entire protein sequence of each antigen. The results (Table 25) are expressed as the number of spot-forming cells (sfc) per million peripheral blood mononuclear cells (PBMC) that respond to each of the peptide pools.

Results indicate the following observations: (1) each of the single gene constructs (MRKAd5gag, MRKAd5pol, or MRKAd5nef) is able to elicit high levels of antigen-specific T cells in monkeys; (2) the single-gene MRKAd5 constructs can be mixed as a multi-cocktail formulation capable of eliciting very broad T cell responses against gag, pol, and nef; (3) the MRKAd5 vector expressing the fusion

protein of gag plus IA pol is capable of inducing strong T cell responses to both gag and pol.

Table 25. Evaluation of Mixtures of MRKAd5 vectors expressing humanized

5 HIV-1 gag, pol, gagpol, nef in rhesus macaques

10

| Grp# | Vaccine                       | Monk # |      |       | T=6 wks |         |      |
|------|-------------------------------|--------|------|-------|---------|---------|------|
| •    | T=0, 4 wks                    |        | Mock | Gag H | Pol - 1 | Pol - 2 | Nef  |
| 1    | MRKAd5 gag                    | CB9V   | 0    | 15    | •       | -       | -    |
|      | 10^10 vp                      | CD19   | 0 .  | 374   |         | •       | -    |
| ļ    |                               | 109H   | 1    | 843   | -       | •       | -    |
| 2    | MRKAd5 gag                    | 99D130 | 1    | 948   | •       | •       |      |
|      | 10^8 vp                       | W277   | 16   | 324   | •       | •       | -    |
|      |                               | 143H   | 4    | 595   | -       | -       | •    |
| 3    | MRKAd5 poi                    | CC1X   | 4    | •     | 46      | 256     | -    |
| _    | 10^10 vp                      | AW3W   | 3    | -     | 463     | 550     | -    |
| ĺ    |                               | AV43   | 6    | -     | 95      | 1333    | -    |
| 4    | MRKAd5 pol                    | AW38   | 1    | -     | 19      | 30      | -    |
| 1    | 10^8 vp                       | CC8K   | 0    |       | 50      | 995     | ٠ ا  |
|      |                               | CC21   | 1    | -     | 33      | 436     | -    |
| 5    | MRKAd5 nef                    | 076Q   | 9    | -     | -       | -       | 1204 |
| - 1  | 10^10 vp                      | 091Q   | 4    | -     | -       | ٠ .     | 85   |
|      |                               | 083Q   | 0    | -     | -       | -       | 176  |
| 6    | MRKAd5 nef                    | 000029 | 1    | -     | -       | -       | 114  |
|      | 10^8 vp                       | 98D022 | 6    |       | -       | -       | 170  |
| 1    | •                             | 98D160 | 3    | -     | -       | -       | 198  |
| 7    | MRKAd5gag+MRKAd5pol+MRKAd5nef | 99D251 | 3    | 206   | 15      | 193     | 120  |
|      | 10^10 vp each                 | 05H    | 3    | 135   | 21      | 9       | 638  |
|      |                               | 00C016 | 3    | 26    | 4       | 51      | 23   |
| 8    | MRKAd5gag+MRKAd5pol+MRKAd5nef | 99D215 | 1    | 171   | 18      | 193     | 240  |
|      | 10^8 vp each                  | 81H    | 5    | 73    | 6       | 14      | 243  |
|      |                               | 12H    | 8    | 1140  | 115     | 811     | 719  |
| 9    | MRKAd5gagpol +MRKAd5 nef      | 99D211 | 0    | 83    | 56      | 838     | 725  |
|      | 10^10 vp each                 | 22H    | 4    | 385   | 119     | 1194    | 1915 |
|      |                               | 61H    | 4    | 343   | 11      | 765     | 853  |
| 10   | MRKAd5gagpol +MRKAd5 nef      | 34H    | 3    | 78    | 19      | 5       | 75   |
|      | 10^8 vp each                  | 48H    | 1    | 65    | 105     | 46      | 43   |
|      |                               | 70H    | 5    | 158   | 15      | 220     | 191  |

Indicated are numbers of spot-forming cells per million PBMCS against the peptide pools. Mock, no peptides; gag H, fifty 20-aa peptides encompassing p55 sequence; pol-1, 20-aa peptides representing N-terminal half of IA pol; pol-2, 20-aa peptides representing the carboxy-terminal half of IA pol; nef, 20-aa peptides encompassing the entire wild-type nef sequence. Responses to the antigens prior to the first immunization did not exceed 40 sfc/10^6 PBMC.

#### WHAT IS CLAIMED IS

5

10

A recombinant adenoviral vaccine vector at least partially deleted in
 E1 and devoid of E1 activity, comprising:

- a) an adenovirus cis-acting packaging region corresponding to from about base pair 1 to between from about base pair 400 to about base pair 458 of a wildtype adenovirus genome; and
- b) a gene encoding an HIV protein or immunologically relevant modification thereof.
- A vector in accordance with claim 1 comprising a packaging region corresponding to from about base pair 1 to about base pair 450 of a wildtype adenovirus genome.
- 3. A vector in accordance with claim 1 further comprising nucleotides
   15 corresponding to between from about base pair 3511 to about 3524 to about base pair
   5798 of a wildtype adenovirus genome.
  - 4. A vector in accordance with claim 3 comprising base pairs corresponding to 1-450 and 3511-5798 of a wildtype adenovirus genome.
- 5. A vector in accordance with claim 4 which is deleted of base pairs20 451-3510.
  - 6. A vector in accordance with claim 1 which is at least partially deleted in E3.
  - 7. A vector in accordance with claim 6 wherein the E3 deleted region is from base pairs 28,133-30,818.

8. A vector in accordance with claim 1 wherein the gene encoding the HIV protein or modification thereof comprises codons optimized for expression in a human.

- 9. A vector in accordance with claim 1 wherein the vector comprises a gene expression cassette comprising:
  - a) a nucleic acid encoding a protein;
  - b) a heterologous promoter operatively linked to the nucleic acid encoding the protein; and
    - (c) a transcription termination sequence.
- 10. A vector in accordance with claim 9 wherein the gene expression cassette is inserted into the E1 region.
  - 11. An adenoviral vector in accordance with claim 9 wherein the gene expression cassette is in an E1 parallel orientation
- 12. An adenoviral vector in accordance with claim 9 wherein the geneexpression cassette is in an E1 antiparallel orientation.
  - 13. An adenoviral vector in accordance with claim 9 wherein the promoter is a cytomegalovirus promoter devoid of intronic sequences.
  - 14. An adenoviral vector in accordance with claim 13 wherein the promoter is an immediate early human cytomegalovirus promoter.
- 20 15. An adenoviral vector in accordance with claim 9 wherein the promoter is a murine cytomegalovirus promoter.
  - 16. An adenoviral vector in accordance with claim 9 wherein the transcription termination sequence is a bovine growth hormone polyadenylation and transcription termination sequence.

17. An adenoviral vector in accordance with claim 9 wherein the transcription termination sequence is a synthetic polyadenylation signal (SPA).

- 18. A cell comprising the adenoviral vector of claim 1.
- 19. Recombinant, replication-defective adenovirus particles harvested
   and purified subsequent to transfection of the adenoviral vector of claim 1 into a cell
   line which expresses adenovirus E1 protein at complementing levels.
  - 20. An HIV vaccine composition comprising purified adenovirus particles of claim 19.
- 21. An HTV vaccine composition of claim 20 which comprises aphysiologically acceptable carrier.
  - 22. A method of producing recombinant, replication defective adenovirus particles containing the adenoviral genome of the adenoviral vector of claim 1 which comprises introducing the adenoviral vector into a host cell which expresses adenoviral E1 protein, and harvesting the resultant recombinant, replication-defective adenovirus.
  - 23. A method according to claim 22 wherein the cell is a PER.C6® cell.

- 24. A method of generating a cellular-mediated immune response
   against HIV in an individual comprising administering to the individual a vaccine of
   claim 21.
  - 25. A method according to claim 24 which further comprises administration to the individual a DNA plasmid vaccine, optionally administered with a biologically effective adjuvant, protein or other agent capable of increasing the immune response.

26. A method according to claim 25 wherein the DNA plasmid vaccine is administered to the individual prior to administration of an adenovirus vaccine.

- 27. A method according to claim 24 wherein the adenovirus vaccine is
   5 preceded by an adenovirus vaccine of a different serotype.
  - 28. A method according to claim 24 which comprises administering and readministering the adenovirus vaccine vector to the individual.
  - 29. An adenoviral vector in accordance with claim 1 wherein the HIV protein is HIV gag or an immunologically relevant modification thereof.
- 30. An adenoviral vector in accordance with claim 9 wherein the gene expression cassette comprises an open reading frame encoding an HIV gag protein or immunologically relevant modification thereof.
  - 31. A recombinant adenoviral vaccine vector at least partially deleted in E1 and devoid of E1 activity, comprising:
- a) an adenovirus cis-acting packaging region corresponding to from about base pair 1 to about base pair 450 and from about 3511 to about 5798 of a wildtype adenovirus genome, and deleted for base pairs corresponding to from about base pair 451 to from about base pair 3510 of a wildtype adenovirus genome; and
  - b) a gene expression cassette comprising

- i) SEQ ID NO: 29;
- ii) a heterologous promoter operatively linked to i); and
- iii) a transcription termination sequence.

32. An adenoviral vector in accordance with claim 31 wherein the gene expression cassette is in an E1 parallel orientation.

- 33 An adenoviral vector in accordance with claim 31 wherein the gene expression cassette is in an E1 antiparallel orientation.
- 34. An adenoviral vector in accordance with claim 31 wherein the promoter is a cytomegalovirus promoter devoid of intronic sequences.

5

- 35. An adenoviral vector in accordance with claim 31 wherein the transcription termination sequence is a bovine growth hormone polyadenylation and transcription termination sequence.
- 36. An adenoviral vector in accordance with claim 31 which is at least partially deleted in E3.
  - 37. A cell comprising the adenoviral vector of claim 30.
  - 38. Recombinant, replication-defective adenovirus particles harvested and purified subsequent to transfection of the adenoviral vector of claim 30 into a cell line which expresses adenovirus E1 protein at complementing levels.
  - 39. An HIV vaccine composition comprising purified adenovirus particles of claim 38.
  - 40. An HIV vaccine composition of claim 39 which comprises a physiologically acceptable carrier.
- 41. A method of producing recombinant, replication defective adenovirus particles containing the adenoviral genome of the adenoviral vector of claim 30 which comprises introducing the adenoviral vector into a host cell which expresses adenoviral E1 protein, and harvesting the resultant recombinant, replication-defective adenovirus.

42. A method according to claim 41 wherein the cell is a PER.C6® cell.

43. A method of generating a cellular-mediated immune response against HIV in an individual comprising administering to the individual a vaccine of claim 21.

5

- 44. A method according to claim 43 which further comprises administration to the individual a DNA plasmid vaccine, optionally administered with a biologically effective adjuvant, protein or other agent capable of increasing the immune response.
- 45. A method according to claim 44 wherein the DNA plasmid vaccine is administered to the individual prior to administration of an adenovirus vaccine.
  - 46. A method according to claim 43 wherein the adenovirus vaccine is preceded by an adenovirus vaccine of a different serotype.
  - 47. A method according to claim 43 which comprises administering and readministering the adenovirus vaccine vector to the individual.
  - 48. An adenoviral vector in accordance with claim 1 wherein the HIV protein is HIV pol or an immunologically relevant modification thereof.
- 49. An adenoviral vector in accordance with claim 9 wherein the gene
   20 expression cassette comprises an open reading frame encoding an HIV pol protein or immunologically relevant modification thereof.
  - 50. A recombinant adenoviral vaccine vector at least partially deleted in E1 and devoid of E1 activity, comprising:

a) an adenovirus cis-acting packaging region corresponding to from about base pair 1 to about base pair 450 and from about 3511 to about 5798 of a wildtype adenovirus genome, and deleted for base pairs corresponding to from about base pair 451 to from about base pair 3510 of a wildtype adenovirus genome; and

5

10

- b) a gene expression cassette comprising
  - a nucleotide sequence selected the group consisting of SEQ ID NO: 1, SEQ ID NO: 5 and SEQ ID NO: 7;
  - ii) a heterologous promoter operatively linked to i); and
  - iii) a transcription termination sequence.
- 51. An adenoviral vector in accordance with claim 50 wherein the gene expression cassette is in an E1 parallel orientation.
- 52. An adenoviral vector in accordance with claim 50 wherein the gene expression cassette is in an E1 antiparallel orientation.
- 53. An adenoviral vector in accordance with claim 50 wherein the promoter is a cytomegalovirus promoter devoid of intronic sequences.
- 54. An adenoviral vector in accordance with claim 50 wherein the transcription termination sequence is a bovine growth hormone polyadenylation and transcription termination sequence.
- 55. An adenoviral vector in accordance with claim 50 which is at least partially deleted in E3.
  - 56. A cell comprising the adenoviral vector of claim 49.

57. Recombinant, replication-defective adenovirus particles harvested and purified subsequent to transfection of the adenoviral vector of claim 49 into a cell line which expresses adenovirus E1 protein at complementing levels.

- 58. An HIV vaccine composition comprising purified adenovirusparticles of claim 57.
  - 59. An HIV vaccine composition of claim 58 which comprises a physiologically acceptable carrier.
  - 60. A method of producing recombinant, replication defective adenovirus particles containing the adenoviral genome of the adenoviral vector of claim 49 which comprises introducing the adenoviral vector into a host cell which expresses adenoviral E1 protein, and harvesting the resultant recombinant, replication-defective adenovirus.

10

- 61. A method according to claim 60 wherein the cell is a PER.C6® cell.
- 62. A method of generating a cellular-mediated immune response against HIV in an individual comprising administering to the individual a vaccine of claim 59.
  - 63. A method according to claim 62 which further comprises administration to the individual a DNA plasmid vaccine, optionally administered with a biologically effective adjuvant, protein or other agent capable of increasing the immune response.
  - 64. A method according to claim 63 wherein the DNA plasmid vaccine is administered to the individual prior to administration of an adenovirus vaccine.

65. A method according to claim 62 wherein the adenovirus vaccine is preceded by an adenovirus vaccine of a different serotype.

- 66. A method according to claim 62 which comprises administering and readministering the adenovirus vaccine vector to the individual.
- 67. An adenoviral vector in accordance with claim 1 wherein the HIV protein is HIV nef or an immunologically relevant modification thereof.

5

15

- 68. An adenoviral vector in accordance with claim 9 wherein the gene expression cassette comprises an open reading frame encoding an HIV nef protein or immunologically relevant modification thereof.
- 10 69. A recombinant adenoviral vaccine vector at least partially deleted in E1 and devoid of E1 activity, comprising:
  - a) an adenovirus cis-acting packaging region corresponding to from about base pair 1 to about base pair 450 and from about 3511 to about 5798 of a wildtype adenovirus genome, and deleted for base pairs corresponding to from about base pair 451 to from about base pair 3510 of a wildtype adenovirus genome; and
  - b) a gene expression cassette comprising
    - a nucleotide sequence selected the group consisting of SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13 and SEQ ID NO: 15;
    - ii) a heterologous promoter operatively linked to i); and
    - iii) a transcription termination sequence.
  - 70. An adenoviral vector in accordance with claim 69 wherein the gene expression cassette is in an E1 parallel orientation.

71. An adenoviral vector in accordance with claim 69 wherein the gene expression cassette is in an E1 antiparallel orientation.

- 72. An adenoviral vector in accordance with claim 69 wherein the promoter is a cytomegalovirus promoter devoid of intronic sequences.
- 73. An adenoviral vector in accordance with claim 69 wherein the transcription termination sequence is a bovine growth hormone polyadenylation and transcription termination sequence.
  - 74. An adenoviral vector in accordance with claim 69 which is at least partially deleted in E3.
    - 75. A cell comprising the adenoviral vector of claim 68.

10

- 76. Recombinant, replication-defective adenovirus particles harvested and purified subsequent to transfection of the adenoviral vector of claim 68 into a cell line which expresses adenovirus E1 protein at complementing levels.
- 77. An HIV vaccine composition comprising purified adenovirusparticles of claim 76.
  - 78. An HIV vaccine composition of claim 77 which comprises a physiologically acceptable carrier.
  - 79. A method of producing recombinant, replication defective adenovirus particles containing the adenoviral genome of the adenoviral vector of claim 68 which comprises introducing the adenoviral vector into a host cell which expresses adenoviral E1 protein, and harvesting the resultant recombinant, replication-defective adenovirus.
  - 80. A method according to claim 79 wherein the cell is a PER.C6® cell.

81. A method of generating a cellular-mediated immune response against HIV in an individual comprising administering to the individual a vaccine of claim 78.

- 82. A method according to claim 81 which further comprises

  administration to the individual a DNA plasmid vaccine, optionally administered with
  a biologically effective adjuvant, protein or other agent capable of increasing the
  immune response.
- 83. A method according to claim 82 wherein the DNA plasmid
   vaccine is administered to the individual prior to administration of an adenovirus
   vaccine.
  - 84. A method according to claim 81 wherein the adenovirus vaccine is preceded by an adenovirus vaccine of a different serotype.
  - 85. A method according to claim 81 which comprises administering and readministering the adenovirus vaccine vector to the individual.
- recombinant, replication-defective adenovirus particles, wherein the adenovirus particles are harvested and purified from a cell line expressing adenovirus E1 protein, and wherein the particles are harvested subsequent to transfection of the cells with an adenoviral vector or vectors in accordance with claim 9; said vector(s) comprising a gene expression cassette or cassettes comprising nucleotide sequences encoding HIV proteins selected from the group consisting of:
  - a) gag, pol, and nef, expressed independently from three individual vectors;

 b) gag, pol, and nef, expressed independently from one vector with the encoding nucleic acid sequences operatively linked to distinct promoters and transcription termination sequences;

- c) gag, pol, and nef, expressed via two vectors, one expressing a polnef fusion, and another expressing gag;
- d) gag, pol, and nef, expressed via two vectors, one expressing a gagpol fusion and another expressing nef;
- e) gag, pol and nef, expressed via two vectors, one expressing a nefgag fusion and another expressing pol;
- f) gag, pol, and nef, expressed via one vector expressing a gag-polnef fusion;
- g) gag and pol, expressed independently from two individual vectors;
- h) gag and pol, expressed independently from one vector with the encoding nucleic acid sequences operatively linked to distinct promoters and transcription termination sequences;
- i) pol and nef, expressed independently from two individual vectors;
- j) pol and nef, expressed independently from one vector with the encoding nucleic acid sequences operatively linked to distinct promoters and transcription termination sequences;
- k) nef and gag, expressed independently from two individual vectors;
- nef and gag, expressed independently from one vector with the encoding nucleic acid sequences operatively linked to distinct promoters and transcription termination sequences;
- m) gag and pol, expressed via one vector expressing a gag-pol fusion;

5

10

15

n) pol and nef, expressed via one vector expressing a pol-nef fusion;
 and

- o) nef and gag, expressed via one vector expressing a nef-gag fusion.
- 87. A multivalent adenovirus vaccine composition in accordance with claim 86 wherein the gag-pol fusion consists of SEQ ID NO: 39.
  - 88. A multivalent adenovirus vaccine composition in accordance with claim 86 wherein the fused sequences have the encoding nucleic acid sequences operatively linked to distinct promoters and transcription termination sequences.
- 89. A multivalent adenovirus vaccine composition in accordance with

  10 claim 86 wherein the fused sequences have the encoding nucleic acid sequences

  operatively linked to a single promoter; and the encoding nucleic acid sequences

  operatively linked by an internal ribosome entry sequence ("IRES").

# Original Adenovector Construct:

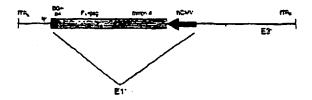


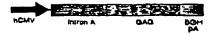
Figure 1: Original HIV-1 gag adenovector.

# Sequence of the open reading frame for FL-qaq (human codon optimized)

atgggtgctagggcttctgtgctgtctggtggtgagctggacaagtgggagaagatcaggctgaggcctggtgg caagaagaagtacaagctaaagcacattglgtgggcctccagggagctggagaggtttgctgtgaaccctggc cigciggagacticigagggigcaggcagatecigggcagctccagccctcccigcaaacaggctcigagg agcigaggicccigiacaacacagiggciacccigiacigigigcaccagaagatigatgigaaggacaccaag gaggecciggagaagatigaggaggagcagaacaagtccaagaagaaggcccagcaggctgctgctggc acaggcaactccagccaggtgtcccagaactaccccattgtgcagaacctccagggccagatggtgcaccag gccatctcccccggaccctgaatgcctgggtgaaggtggaggaggaggaggcttctcccctgaggtgatccc catgitctctgccctgtctgagggtgccacccccaggacctgaacaccatgctgaacacagtgggggggccatc aggcigccatgcagatgctgaaggagaccatcaatgaggaggctgctgagtgggacaggctgcatcctgtgc acgciggccccattgccccggccagaigagggagcccaggggctctgacattgctggcaccacctccaccct ccaggagcagattggctggatgaccaaccaccccccatccctgtgggggaaatctacaagaggtggatcat ccttcagggactatgtggacaggttctacaagaccctgagggctgagcaggcctcccaggaggtgaagaact ggatgacagagaccctgctggtgcagaatgccaaccctgactgcaagaccatcctgaaggccctgggccctg gctgaggccatgtcccaggtgaccaactccgccaccatcatgatgcagaggggcaacttcaggaaccagag gaagacagtgaagtgcttcaactgtggcaaggtgggccacattgccaagaactgtagggcccccaggaaga cccaaqaqiccticaggtttgggaggagaagaccaccccagccagaagcaggagcccattgacaagg ageigiaeccetiggeticetigaggicetigttiggeaaegacceteticecaglaaaalaaageeegggea gat (SEQ ID NO: 29)

Figure 2

## Old Transgene:



# New Transgenes:

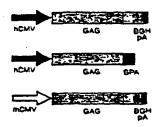


Figure 3: Diagrammatic representation of the original HIV-1 gag transgene and the series of new transgene constructions.

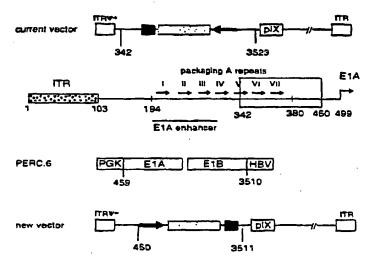


Figure 4: Modifications made to the current adenovector backbone in the generation of the new vector.

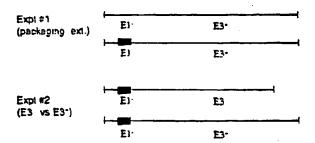


Figure 5: Virus mixing experiments to determine the effects of the addition made to the packaging signal region (Expt #1) and analysis of the effects of the E3 gene on viral growth (Expt. #2). The red bars denote the region of modifications made to the E1 deletion.



Figure 6: Autoradiograph of viral DNA analysis following viral mixing experiments (expts. #1 and #2) as detailed in the text.

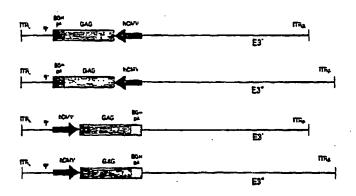


Figure 7A: hCMV-FLgag-bGHpA adenovectors constructed within the "MRK" backbone. E1 parallel and E1 antiparallel transgene orientation within the E3- and E3+ backbones were constructed.

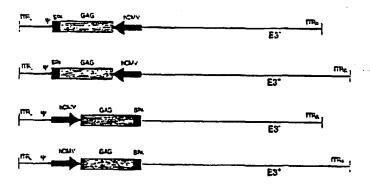


Figure 7B: hCMV-FLgag-SPA adenovectors constructed within the "MRK" backbone. E1 parallel and E1 antiparallel transgene orientation within the E3- and E3+ backbones were constructed.

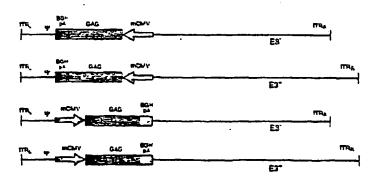


Figure 7C: mCMV-FLgag-bGHpA adenovectors constructed within the \*MRK\* backbone. E1 parallel and E1 antiparallel transgene orientation within the E3- and E3+ backbones were constructed.

WO 02/022080

# Plasmid mixing expt: (orientation)

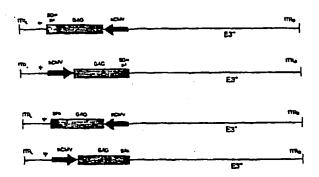


Figure 8A: Effect of transgene orientation

#### Plasmid Mixing expt: (poly A signal)

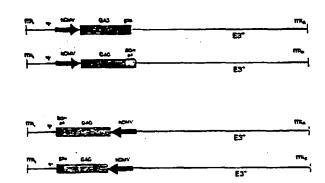


Figure 8B: Effect of polyadenylation signal

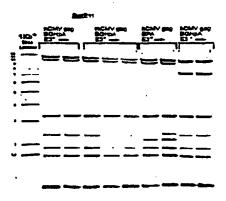


Figure 9: Viral DNA from the four Adgag candidates at P5, following BsfE11 digestion.

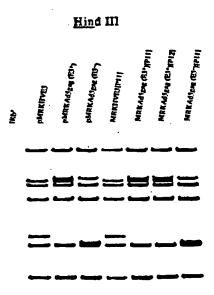


Figure 10: Viral DNA analysis of passage 11 and/or 12 of MRKHVE3, MRKAd5gag and MRKAd5gag(E3-).

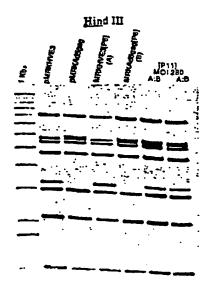


Figure 11: Viral DNA analysis (*Hind*IIII digestion) of passage 6 MRKHVE3 and MRKAd5gag used to initiate the viral competition study. Last two lanes are passage 11 analysis of duplicate passages of the competition study (each virus at MOI 280 vp).

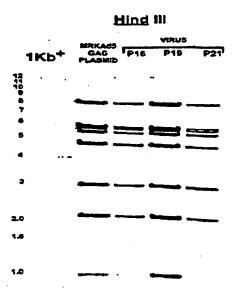
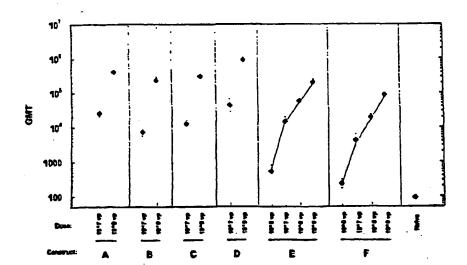


Figure 12: Viral DNA analysis by HindIII digestion on high passage numbers for MRKAd5gag in serum containing media with collections made at specified times. The first lane shows the 1 Kb DNA size marker. The other lanes represent pre-plasmid control (digested with Pac1 and HindIII), and MRKAd5gag virus continually passaged to P16, P19 and P21(serum containing media).

Figure . Serum anti-p24 Levels at 3 Wks post i.m. immunization of balb's mice (n=10) with Varying Doses of Several Adgag constructs: (A) MRKAd5gag (through passage 5): (B) MRKAd5 E3 bCMV-FLgag-bGHpA; (C) MRKAd5 E3 bCMV-FLgag-SPA; (D) MRKAd5 E3 mCMV-FLgag-bGHpA; (D) research Lot (293 cell-derived) of Ad5HIV-1gag; and (F) clinical lot (Ad5gagFN0001) of Ad5HIV-1gag. Reponde are the geometric mean titers (GMT) for each cohort.



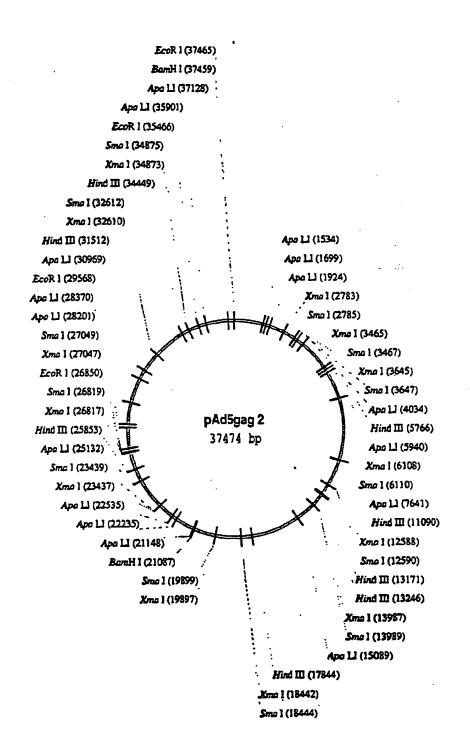


Figure 14

| •           |                   | ŧ            |               |                     | -                  | A STATE OF THE PARTY IN  |              | STATE OF THE PARTY |              | CTCCCAACCC           |
|-------------|-------------------|--------------|---------------|---------------------|--------------------|--------------------------|--------------|--|--------------|----------------------|
| _           | · TTCTTANTTA      | ACATCATCA    | TANTATALLT    | TATTITION           |                    | אייין איין איין איין איי | 2000         |  |              |                      |
|             | ANDANTTAAT        | TISTACTACTT  | ATTATATTA     | ATANAACCTA          | אידוי יידוא        | TACTATTACT COCCACCTO     | ניכבכעניכוב  | AACALTISCA   | רנינייינייני | בארכו בוניי          |
| 101         | CACCOCCTOAC       | GTAGTAGTGT   | GREGATANTE    | TriATriTTYICA       | ACTITION           | AACACATGEA               | ACKINACCIONT | CTCCCANANG   | TOACCOUNTY   | ממוגשתאנט.           |
|             | CCCCCCACTG        |              | CCCCCTTCAC    | ACTACAACGT          | <b>TUNCACOCOC</b>  | TTGTGTACAT               | TOSCIBIOGIA  | CACCOTTATIC  | ACTOCANANA   | CCACACACA            |
| 201         | CATHETACACA       | -            | ATTENTACE     | GETTTTACCC          | GGATOTTGTA         | CIAMITICS                | CATCITANCEGA | GTAAGATTTG   | OCCATITITION | CHECKANA             |
| 1           | CCACATOTOT        |              | TAMARICAN     | CCANANTCCO          | CCTACAACAT         | CATTITAMOC               | COCATTORICT  | CATTCTAAAC   | COCTANAAGC   | (ACCC1TFINA)         |
| 101         | CANTABORD         |              | TCAATAATT     | Trittitud           | ATMYTHYTH          | ATATTTGICT               | ACKICCCATGG  | GGACTITIGAC  | CONTINCEN    | GASACTICTA TO        |
| !<br>?<br>} | CTTATTCTCC        |              | ACTIVITAM     | ACACAAATGAG         | TATEMICHICAT       | TATAAACAAA               | Tecendedec   | CCTCAAACTO   | GCAANTGCAC   | CTCTGAGCO.           |
| 100         | CACATICIPIT       | Perchaggier  | THECKICAL     | CCOCOTCAM           | CHICKNOCHTT        | TATTATTATA               | סטטטטטממטט   | ATCCATTGCA   | TACGETGEAT   | CCATATEAT            |
|             | GTCCACAAAA        |              | MAKKKICAA     | GCCCAGTTF           | CANCYCICANA        | ATAATAATAT               | ככמככמסכמכ   | TAGGTAACGT   | ATCCAACATA   | CCTATAGTAT.          |
| 501         | ATATOTACAT        | TTATATIOGC   | TCATGTCCAA    | CATTACCGCC          | ATCITICACAT        | TCATTATTCA               | CTACTTATTA   | ATAGTAATCA   | ATTACGGGGT   | CATTAGETICA          |
|             | TATACATOTA        | ANTATARCCO   | AGTACAGGIT    | CTAATOCCCG          | TACAACTGTA         | ACTAATAACT               | GATCAATAAT   | TATCATTAGE   | TAATGCCCCA   | GTANTCAMGT           |
| 601         | TACCCATAT         | ATGGAGTTCC   | GCOTTACATA    | ACTITACIONIA        | ANTRACCIONC        | CTYNOCTONOCC             | שכככישכניים  | CCCCOCCCAT   | TOACGICAAT   | AATGACGTAT           |
|             | ATCCCCTATA        |              | COCMATGTAT    | TCAATCCCAT          | TTACCGGGG          | CACCCACTEG               | CGGGTTGCTG   | GROCCOCOTA   | ACTOCAGTTA   | TTACTOCATA           |
| 101         | GITCCCATAG        | TAACGCCAAT   | AGGGACTITIC   | CATTRIACRITE        | ANTCAGATAGA        | GTATITIACIA              | TANACTOCC    | ACTIGOCAUT   | ACATCANOTO   | TATCATATES           |
| }           | CANOGOTATO        |              | TCCCTOAAA     | <b>OTANCTICION</b>  | TTACCCARRET        | CATANATISIC              | ATTIGACOO    | TOWACCOTICA  | TOTAOTICAC   | ATAGTATACC           |
| 108         | CAACTPACKIC       | _            | GTCAATGACG    | CTAAATTCCC          | CCCCTTAGGAT        | TATGCCCAGT               | ACATGACCTT   | ATCCCACTIT   | CCTACTTOGC   | AGTACATOTA           |
| !           | GITCATGCGG        | _            | CNOTTACTOC    | CATTTACCOS          | CCCGACCGTA         | ATACCCCTTCA              | TGTACTCGAA   | TACCCTGANA   | GCATCAACCO   | TCATGTAGAT           |
| 106         | CATATTAGE         | -            | CCATCCTCAT    | CCCCTTTTCC          | CASTACATCA         | ATISGRACATIO             | ATAGGGGTTT   | CACTCACGGG   | DATTTCCAAD   | TCTCACCC.            |
|             | CCATAATCAG        |              | COTACCACTA    |                     | <b>OTCATVITAGE</b> | TACCCCCACC               | TATCOCCANA   | CHEMOTOCCC   | CTAAAGGTTC   | AGAGGTAGG            |
| 1001        | ATTOMOTOR         |              | GFFFFGGCAC    | CAMATCANC           | CAGACTITICS        | ANATISTICGE              | AACAACTCCO   | CCCCATTGAC   | GCAAATGGGC   | GOTAN/2COTI1         |
|             | TAACTOCAGE        |              | CAAAACCGTO    | GTTTTAGTTG          | CCCTGAAAGG         | TITTACAGCA               | THETTOROGE   | COCOTANCTO   | COTITACCCO   | CCATICKECAC          |
| 1101        | TACOOTOOOA        | _            | ACCAGAGETE    | GTTTAGTGAA          | CCGTCAGATC         | GCCTATAGAC               | GCCATCCACG   | CTGTTTTGAC   | CTCCATAGAA   | מאכאהכמההא           |
|             | ATOCCACCCT        | CCAGATATAT   | TCGTCTCGAG    | CANATICACTT         | CCCAGTCTAG         | COGNICACIO               | COGTARGIBC   | GNCANANCTO   | CAGGTATCTT   | Crordactur           |
|             |                   |              |               |                     |                    |                          | Rgill        |  |              |                      |
| 1201        | CTURTICAGE        | CHECOCOGCE   | GOGAACOOTO    | CATTOGRACE          | COGNITICOCC        | GTYZCTAARAG              | TRAGATICTAC  | CATGGGTGCT   | AGGCTTCTG    | TOCTOTOTOG           |
|             | COCTACCTCC        | -            | CCCTTGCCAC    | CITACCTICC          | GCCTAARAGG         | CACCANTACTO              | ACTCTAGATO   | GTACCCACGA   | TCCCGANGAC   | ALGAL MONCE          |
| 1301        | TYCHOACTO         |              | AGAAGATCAG    | <b>GCTRIABORICT</b> | COLUMBIA           | ACANCTACAA               | CCTAAAGCAC   | ATTOTOTOGO   | CCTCCAGGGA   | בעבור זוין אויין פון |
|             | ACCACTCGAC        |              | TCTTCTAGTC    | CCACTCCCCA          | CCACCGFTCT         | TUTTICARGET              | CGATTTCGTG   | TANCACACC  | COMPACTOR    | בפתררורויר           |
| 1401        | Triocretch        | ACCCTODOCT   | GCTOGAGACC    |                     | CCAGCCAGAT         | CCTGGGGCAG               | CHCCAGCCC    | CCCTOCAAAC   | ACCEPTIONS   | CAGCTGAGGT           |
|             | AAACGACACT        | TOGGACCOGA   | COACCTCTOO    | AGACTCCCCA          | COTCOCTCTA         | מבשכככנמנוב              | GAGGICGGGA   | ואסשרטווופ   | ורכושמערור   | 44.744.000           |
| 1501        | CCCTOTACAA        | -            | ACCENTACE     |                     | GANCATTICAT        | CITTAMERIALA             | CUANGRACIC   | CCTCGAGAAG   | TARCTICATOR  | TCOTCTTOFT           |
|             | OCCACATOTT        | P STOTCACCOA | TOGOACATGA    |                     | CTTTTARCTA         | CM.TRC IG                | פיוורורים    | Section 1  |              | CCAGATANTE           |
| 1601        | <b>GTCCAARANO</b> |              | ARREPRETER    |                     | AACTICACK          | ACTACHECEA<br>TOTAL BOOK | CANCIDACET.  | TAACACGTCT   | TOGAGGICCC   | GOTCTACCAC           |
|             | CADOTICITO        | : Treconstra | TCCCGACCGACCG | ACCURCACTO          | יאו יוויאאיירד     | TI, CM, MIRAN            |              |  |              |                      |

Figure ISA

# PMRKANISAAA MER682

| 1701 | CACCAGGCCA                              | TCTCCCCCC          | CACCCTCAAT           | CCCTCTCTCTCA        | ACCEPTATIONS    | CAMANANGA                   | THYTECTECTS       | ACCIDENTECE                             | CATOTICICI   | GCCCTGTCTC       |
|------|---|--------------------|----------------------|---------------------|-----------------|-----------------------------|-------------------|---|--|------------------|
| 1001 | ACCONTOCCAC                             | CCCCCARROAC        | CTCAACACTA           | Trit.Triyyr.YC      | ACTIVITY        | CA'IY'ACKKITTE              | CCATCX ACAT       | GCTGAAGGAG                              | ACCATCAATO   | ACKINGCTTS       |
|      | TCCCACGGTG                              | <b>GOOGOTICCTO</b> | GACTIVITIAN          | ACTIAL TATAGE       | TCN 11 TCCTG    | מדאחדוניהמאני               | COTACCITCIA       | CCACTICCTC                              | TOGTACTTAC   | TCCTTCCSAL: 1    |
| 1901 | TOAOTGGGAC                              | ACCCTCCATC         | CTCTMCMXXC           | TRACICCATT          | שמעמנונונייייי  | אטטטעיאנאטע                 | CACCAGGAAAC       | <b>PCTGACATTG</b>                       | CTORCACCAC   | בובכניוטניבוי.   |
|      | ACTCACCCTO                              | TCCOACGTAG         | GACACTETICA          | AC.CCRRRITAA        | מטענאאטט        | T.TACTOCT                   | כנוטטובנינים      | AGACTOTAAC                              | GACCOTOGTO   | GACICITOSCAVE    |
| 2001 | CAGGAGCAGA                              | PROCERCAN          | GACKINATING          | CCCCCATC            | VIXED KALLALD   | ANTICTACAAG                 | NYTHANATCA        | TCCTONCCT                               | GANCARGATT   | OFFINERATE       |
|      | Greencoter                              | MCCGACCTA          | CTCOTTCTTG           | CONTRACTAGE         | מאנאנכנמים      | THAGATETTE                  | TCCACCTAGE        | AGGACCCCGGA                             | CTTCTTCTA  | CACTCCTACA       |
| 2101 | ACTOCOCCORC                             | CTCCATCCTG         | GACATCARAGE          | ACCECTAN            | ואיאיגיניליזאני | ACCONCTATO                  | TROACAGGTT        | CTACAAGACC                              | CTGAGGGCTG   | AGCANGCCT"       |
|      | TOAGGGGGTG                              | DAGGTAGGAC         | CTGTMFTCCG           | recedederi          | CCTCCGATANG     | TCCCTGATAC                  | ACCTUTCCA         | GATGTTCTGG                              | GACTCCCGAC   | TOTALCOGALI      |
| 2201 | CCAGGAGGTO                              | AAGAACTORA         | TOACAGAGAC           | CCTGCTGGTG          | CANAATGEEA      | ACCEPTACTO                  | CAAGACCATC        | CTGAAGGCCC                              | roancertoc   | TOCCALCCTO       |
|      | OGTCCTCCAC                              | TICTIOACCT         | ACTORCTICTO          | GOACHACCAC          | CTCTTACGGT      | TOCHACTICAC                 | CITCTCGTAG        | GACTTCCGGG                              | ACCCCCCONCO  | ACCOTOCON.       |
| 1062 | CACCACATCA                              | TOACACCTO          | CCACCACCAG           | פנאספנינכזמ         | CHICACAAAGC     | CAGAGINETIC                 | GCTGAGACCA        | TOPCCCAGGE                              | GACCAACTCC   | <b>GCCACCATC</b> |
|      | CHCCACTACT                              | ACTOTOGGAC         | OCHCCCCCAC           | CCCCCGGGAC          | CASTRITTICA     | GTCCCACCAC                  | COACTCCCCT        | ACAGGGGTCCA                             | CTCOTTGAGG   | COCHOCTAGT       |
| 2401 | TGATGCAGAG                              | GOGCAACTIC         | ACCIMACCAGA          | GRANGACAGT          | GAAGTGCTTC      | AACTISTICKA                 | ACCITCOCCCA       | CATTGCCAAG                              | <b>NACTOTACOG</b>  | CCCCCCMOUN.      |
|      | ACTACOTOTO                              | CCCOFFGAAG         | <b>TCCTTGGTCT</b>    | CCTTCTGTCA          | CTTCACGAAG      | TRIACACCUST                 | TCCACCCAGT        | GTAACGGTTC                              | TTGACATCCC   | OCCUPATION !     |
| 2501 | GAAGGCCTCC                              | TOGANOTOTO         | <b>GCANGGANGG</b>    | CCACCAGATO          | ANCHARTOCA      | ATCAGACCA                   | <b>OCCANCITIC</b> | CTGGGGGAAA                              | TCTOOCCCTC   | CCNCNAGGCR:      |
|      | CTTCCCCARCO                             | ACCTTCACAC         | COPPOCINCO           | <b>RETUGITETINE</b> | TICCIRACGE      | TACTUTUCAL                  | CCRATTOMO         | OACCCONTIT                              | AGACCOGGAG   | <b>GOTOTICCU</b> |
| 2601 | AGGCCTGGCA                              | ACTICCTCCA         | GTCCAGGCCT           | <b>DNOCCCUAG</b>    | CCCCTCCCGA      | GOAGTCCTTIC                 | AGOTTTOOR         | ACCACANCAC                              | CACCCCCAGC   | CACIARCCAIT      |
|      | TCCOCACCOT                              | TGAAGGAGGT         | CAGGTCCGGA           | CTCGGGTGTC          | CONSONGRACT     | CCTCAGGANG                  | TCCAMCCCC         | recrement                               | GTGGGGGTCG   | memorie.         |
|      |   |                    |                      |                     |                 |                             |                   |   | -  | II A             |
| 2701 | AGCCCATTOA                              | CAAGGAGCTG         | TACCCCCTTIG          | CCTCCCTOAG          | OTCCCTUTT       | COCAACTACC                  | CCTCCTCCCA        | GTANNITAAA                              | ОСССОВИНСАВ  | ATCTOCTOTA:      |
|      | TCGGGTAACT                              | <b>OFFICETCOAC</b> | ATCCCCCCACC          | הניאממממעדכ         | CAGGGACAAA      | כי יקידיאל דיקה             | COARCARCET        | CATHTTAFFF                              | COOCCCOTC  | TACIACCACA       |
| 2801 | CCTTCTAGT                               | OCCAOCCATC         | TOMOTHEC             | CCCTCCCCCG          | TOCCINCOLL      | DACCTORIAN                  | <b>GGTYCCACTC</b> | CCACTOTCCT                              | TTCCTAATAA   | ANTENCONN        |
| •    | OCHACATCAA                              | COGREGOTAG         | ACANCANACG           | <b>COCACCOCAC</b>   | ACCIGANGGAA     | CTGGGACCTT                  | CCACGGTGAG        | GGTCACAGGA                              | AAGGATTATT   | TIACTOCTT        |
|      |   |                    |                      |                     |                 |                             |                   |   |  | Sphi             |
| 1062 | TTOCATCOCA                              | TTGTCTGAGT         | ADDITION             | CTATTCTORG          | CRASTRAGASTES   | GRECAGEACA                  | CCAACGTTTA        | GCATTOCCAA                              | GACAATAGCA   | GCK:ATCTCRG      |
|      | AACCTAGCGT                              | AACAGACTCA         | TCCACAGTAA           | GATAAGACCC          | CCCACCCCAC      | CCCGTCCTGT                  | CGTTCCCCCT        | CCTANCCCTT                              | CTOTATOOT  | CCCTACGACC       |
|      |   |                    | Pref                 | 7                   |                 |                             |                   |   |  |                  |
| ,    | ,                                       |                    | annanan i            |                     |                 |                             | 10000             |   |  | STORE WILLIAM    |
| 3001 | OCIATOCOOTO                             | COCHCTATOO         | CCGATCGCG CGCCGTACTG | CCCCOTACTG          | AMATCHARA       | CANCELLAR TEN               | ALKAGILIKAJAN     | TOTTATATAT                              |  | GANTACATEN       |
|      |   |                    | מסר זשוררוור         |                     |                 |                             |                   |   |  |                  |
| 3101 | TITOTATCIO                              | TTTTOCAGCA         | BCCGCCGCCG           | CCATriArical        | CAACTECTET      | GATGGAAACA                  | THOMANCHE         | ATATTTCACA                              |  | CCCCATVAGG       |
|      | MANCATAGAC                              | ANACOTOGE          | 2000000000           | COTAC TOTAL         | מאאייאייאייט    | היאביידוניה.<br>היאביידוניה | ANCACTCGAG        | TATAMACTOT                              | TOCOCOTACO   | COCKETACO        |
| 3201 | CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC | CAGAATGTGA         | Trans Trican         | CATTGATTAT          | درردردرزيدر     | _ :                         | CHCTACTACC        | TRACCTACG                               | ACACCCTCTC<br>TOTAL CONTROL  | TCCAACCCCG       |
|      | <b>GCCCCACGCA</b>                       | GICTTACACT         | ACCCTIACAGIC         | GTAACTACCA          | G. Garan. Mai   | W. CA-CO 1 1 1              | CHANGE CONTROL    | 201000000000000000000000000000000000000 | o a construction of the co |                  |

tique 158

# PMRKAd54Ag MER682

| 3301 | TTOORGACTO        |             |               | נא.ויטי.אליזי.אני | ברארימינים                                      |                           | ACTICACTOR                        | CHITCHISAG    | CCCACTIVACA                             | AACAC:TCCAG                 |
|------|-------------------|-------------|---------------|-------------------|---|---------------------------|-----------------------------------|---------------|---|-----------------------------|
| 3401 | CTTCCCOTTC        | ATCENTECTOC | CATCACAAGT    | TCIACCRC TCT      | TTTWEENTAN                                      | TITATATICET               | TOTALCICIANON                     | ACTTAATOTC    | GITTICTCAGE                             | ACKETYCHTYANA               |
|      | GAACCASCAAG       | TAGGCGGCG   | CTACTOTACA    | ACTIVICOGIAGA     | MACCGIVET                                       | MCCTMGM                   | ACTOORICCET                       | TGANTTACAO    | CAANGACTCG                              | TCCACAAM "F                 |
| 3501 | TCTCCCCAO         |             |               |                   | CCCAATTACK                                      | TTTANACAT                 | MANTANAAA                         | CCAGACTETO    | TTIMBATTIG                              | GATCAAGCAA                  |
|      | MORCOCOONC        | GICCAMAGAC  | CCCACTTCCG    | AAGGAAGCAGA       | CERTITACICC                                     | AATITI'R:TA               | TTTATTITE                         | OCTCTCAGAC    | MANCCTANAC                              | CTACITATION I               |
| 3601 | GIGICTIACT        |             |               |                   | ARGCCCCKXGA                                     | מכאמנותיונית              | CONTRACTION                       | CONTRICTOR O  | TATTITICC                               | AGGACGTGET                  |
|      | CACAGAACGA        | CACANATANA  | TCCCCANANC    | GCCCCCCCA         | Treasocet                                       | CONCINCA                  | GCCAGCAACT                        | CCCAGGACAC    | ATAAAAAAGG                              | TCCTGCACCA                  |
| 27   |                   |             |               |                   | ومعادياتها                                      | TOTAL STREET              | Ammuna<br>Acceptating Apprehenden | Anti-Property |   | September 1.                |
| 7    | TTTCCACTOA        |             |               |                   | כאליאניאניכו כ                                  | ACCTICATED                | TECTGACGIC                        | TCCAAGTACG    | ACOCCCCACC                              | ACAACATICTA                 |
| 3801 | DATECABLED        | TACCACCACC  | GCTOGGCGTO    | <b>OTCCCTANA</b>  | ANGRETTICA                                      | GTACKAAGCT                | GATTRICCAGG                       | GOCAGGCCCT    | TOOTOTANOT                              | GTTTACAAN:                  |
|      | CTAGOTCAGC        | ANCONCETED  | COACUTICAC    | CACGGATTIT        | TACAMAGT  | CATCGTTCGA                | CTAACGGTCC                        | CCOTCCGGGA    | ACCACATTCA                              | CAANINGTITIC                |
| 1066 | COUTTABEL         |             |               | GATATOAGAT        | GCATCTTAGA                                      | CAGINATITI                | ACCITAGE TA                       | TUTTCCCAGE    | CATATOCCTC                              | CORRESTITION                |
|      | OCCAMITICAN       |             | _             | CTATACH:TA        | CETW WACC'T                                     | GACATAMAA                 |                                   | ACAMGOGICO    | 200000000000000000000000000000000000000 | פררבי זייים                 |
| 4001 | 1011010CAG        | MCCACCACC   | ACAGTOTATC    | COCHICACTT        | CCCANTITIO                                      | TCAMSTAGET                | TAGAACCAAA                        | ACCENCIONAG   | MACTICOMEN<br>TICANCCIE:                | CCCCCTTCTV                  |
| 4101 | ACCIPITABLE       | •           |               |                   | ATTACAC   | CHARLES CHARLES           | CHADOCGAAG                        |               | GATCACTAAC                              | OTCATAGETO                  |
|      | TOOAGOTTET        | _           |               | TACTACCGT         | TACCOCITIC                                      | נבבעבבעבכם                | OACCCOUNT                         |               | CTAGEGRATIO                             | CANTATCAN                   |
| 4201 | TOTTCCAGGA        | TOAGATEGIC  | ATAXXCATT     | TTTACAAAGC        | GCCACACATAG                                     | COTTACCACAC               | TOCCOTATAL                        | TOGITECATE    | COCCCYGGG                               | GCGTARTTA                   |
|      | ACAAGGTCCT        | ACTCTAGGAG  | TATCCICCTAA   | ANATIGITITES      | COCCUOCUTO                                      | הכאהמפדכים                | ACCCATATE                         | ACCAAGGTAG    | accognice                               | COCATICAN'N:                |
| 4301 | CCTCACACAT        | TROCATTICC  | CACCCTTAGA    | GETCAGATOG        | CHICKLATICATIC                                  |                           |                                   | GAAAACCOSTT   | rccoggang                               | <b>GCGNONTION</b>           |
|      | <b>OGAGTOTOTA</b> | AACCTAAACG  | GTCCGAAACT    | CAAGTCTACC        | CCCCTAGTAC                                      | AGATOGACCC                | CCCCCTACIT                        | CHITTACCAA    | AGGCCCCATC                              | COCTOTAGE                   |
|      |                   |             |               |                   |   |                           |                                   |               |   |                             |
| 4401 | CTCCCAAGAA        | AGCAGGTTCC  | TONOCAGCTO    | CCACTTACCG        | CACICYOGRAG                                     | GCCCGTNANT                |                                   |               | ACTOOTAGET                              | אניאניאניאניא:              |
|      | GACCCTTCTT<br>Pst | TCGTCCAAGG  | ACTCGTCGAC    | OCTOANTOCC        | הדניסטכנאני                                     | COCCUTTIA                 | GTGTGGATAA                        | TOGCCCACOT    | TGACCATCAA                              | FICTOTOGAC                  |
| 4501 | CAGCTGCCGF        | CATCCCTGAG  | CAGGGGGGGCC   |                   | GCATGTCCCT                                      |                           |                                   |               | CAGAAGGCGC                              | TEGECIANCE                  |
|      | GTCGACOOCA        | GTAGGGACTC  | ميدددددوو     | TOMBOMIT          | CGTACAGAIA                                      | CTGAGGGTAC                | ANANGGACT                         | CONTINUENCE   | orcrrccoco                              | AGCGGCGGGT                  |
|      |                   |             |               |                   |   | Difference and the second | Secretaria                        | GACTOTOTOCA   | CLANGCAGT                               | CEAGACACHE                  |
| 4601 | COCTATION         | ANGANCOTTC  | CTTCGTTTCA    | AAAAGTIKGC        | AAACTCTTAIC                                     |                           |                                   |               | GOTTCOTCAA                              | Grinck GCC MI               |
| 4701 | CCACACACTEG       | -           |               | TCGATCCAGE        | ATATOTOCIC                                      | מוווובמכנאכ               | PTCCCCCCCC 1                      |               | CCACCAGTAGT                             | CONTRICTE OFF               |
|      | GOTOTCOAGC        | CAGTGGACGA  | DATCECTING    | ACCTACATICG       | TATAGACAGAG                                     | בששתבנאנכ                 |                                   |               |   | <b>OCCACTOARTY</b>          |
| 4801 | CCAGACGGG         |             | TCTTTCCACG    | CANTO MAKET       | CCTCTTTCAGG                                     | CARCACACC                 | PCACCOTCAA (                      | CAXCHOCOCT    | CCCCGACGC                               | COCTROINCEM.<br>GCCARCOGOTT |
|      | GGICTIGCCCG       | GICCCMFING  | Miscource 11. |                   | W. W. S. W. W. W. W. W. W. W. W. W. W. W. W. W. |                           |                                   |               |   |                             |

figure 15c

# рмпклабаад мек682.

|      | 1                   |                   |                   |                    |                          |                   |               |                   |                   |                        |
|------|---------------------|-------------------|-------------------|--------------------|--------------------------|-------------------|---------------|-------------------|-------------------|------------------------|
| 4901 | <b>CONTROCOCINO</b> | _                 |                   |                    | COGNETICAL               | נונשמישכושכ       | COCCAONTAG    |                   |                   | <b>STECABILITY</b>     |
|      | CCACCCCAAC          | TCCOACCAGO        | ACTIACCACGA       |                    | כדותימכניאתה מכיכאמאתמני | היאכניסרוכאק      | CCTAGINATION  | GTANACTOST        | ACCACAGTAT        | CACIGITECOCI           |
| 2001 | TCCGCGGCGT          | · OCCCCTIVICC     | <b>OCCCARCTTC</b> | CCCTTYXXMX         | AGRICULTURE              | CKINGSTOCKS       | TYCACACTTT    | THACHOCOTA        | CAGCTTVACAC       | <b>OCTIAGAMATIA</b>    |
|      | AGGCGCCGCA          | CCGGGAACCG        | COCOTCOANC        | COGANCUTCE         | Technicat                | CA TRICKEGTE      | ACCITCITIONA  | ACTECEGEAT        | CTCGAACCCO        | COCTUTTAN              |
| 5101 | CCGATTCCCG          | GONCHAGISCA       | TCCCCCCCC         | AGGCCCCCACA        | CALCACTURE               | CATTOCACGA        | CHICKNESSIG   | CTCTGGCCGT        | TCCCCCCCAA        | AMICAGGIT              |
|      | OCCTAABOCC          |                   | -                 |                    | CTRCCAGAGG               |                   | COCHECACIE    | GAGACCOGCA        |                   |                        |
|      |                     |                   |                   |                    |                          |                   |               |                   |                   | (hstt 1071             |
| 5201 | <b>TCCCCCATGC</b>   | TITTERATOR        | GTTTCTTACC        | TETROGETICE        | ATYINGCOOF               | GITCCACGCTC       | GCTCACCAAA    | AGOCTICTOCO       |                   | TOTCCCCOTA TACA!ACTTVI |
|      | AGGGGGTACG          | AANAACTACO        | CAAAGAATOG        | MONCCANAGO         | TACTURE                  | CACACTCCGAG       | CCACTOCTTT    | TCCGACAGGC        | ACAGGGGCAT        | ATCTCTOAM:             |
|      |                     | Khol              |                   | v                  |                          |                   | ÷             |                   |                   | •                      |
| 5301 | <b>AGAGGCCTOF</b>   | U                 | TOTTCCCACGG       | TECTECTER          | ATAGMACTC                | GUACCACTOT        | CACACAAAGG    | CICOCOICCA        | OCCCAGCACO        | AAGGAGGCTA             |
|      | TCTCCGGACA          | OGAGCTCGCC        | ACANGGETACE       | ARCARCACCA         | TATCTTTGAG               | CCTAGGEGAGA       | CTCTOTTTCC    | <b>OAGCOCAGOT</b> | ccoorcoroc        | TYCCTCCGAT             |
| 5401 | ACTOGGAGGG          | OTACCOORCD        | THENCEACTA        | GOCCUTECAC         | Texactecano              | <b>PTGTGAABAC</b> | ACATGTCCCC    | CTCTTCOGCA        | TCAAGGAAGG        | TOATTASTIT             |
|      | TCACCCTCCC          | CATCOCCAGE        | AACACCOTCAT       | CCCCCAGGTG         | NACTARACTIC              | CACACTICITO       | TOTACARCOG    | GAGAAGCCGT        | AGTICCTICC        | ACTANCCANA             |
| 5501 | GTAGGTGTAG          | OCCACOTORC        | COCONCINCO        | TEMOCOCHO          | CTATAAAACG               | GACTOROGIC        | OCCUTOUTE     | TCACTCTT          | CCGCATCGCT        | GTCTGCGAGG             |
|      | CATCCACATC          | COGTOCACTO        | <b>OCCCACAAGO</b> | ACTTCCCCCC         | GATATITICC               | CCCACCCTCG        | CCCAARICAGG   | ACTGAGAGAA        | <b>GOCOTAGOGA</b> | CNGACCETEC             |
| 5601 | OCCAGCION           | COCCICAGIA        | CTCCCTCTOA        | AAAGCGGCA          | TCACTTCTCC               | GCTAAGATTG        | TCAGTTTCCA    | AAAACGAGGA        | COATTTOATA        | TCACCROTY              |
|      | COGTEGACAA          | . CCCCACTCAT      | GAGGGAGACT        | TITCGCCCGF         | ACTGAAGACG               | COATTCTAAC        | ACTCANAGGT    | THEOLOGY          | CCTAMCTAT         | ANOTOGRACCT            |
|      |                     |                   |                   |                    |                          |                   | Himpelli      |                   |                   |                        |
| 5701 | CCGCGCTGAT          | OCCUPATOROD       | <b>ETGCCCCCAT</b> | CCATCTORIC         | ACHANAGACA               | ATCTITION         | _             | OCTOCCAME         | GACCCGTAGA        | CRANCOPTION            |
|      | DOCOCCACTA          | COGNAACTCC        | CACCOCCGTA        | <b>GCTAGACCAG</b>  | TOTATACTOR               | TAGANNAACA        | ACAGITICGNA   | CCACCOTTTG        | CHOOCCATCT        | CCCCCAACCT             |
|      |                     |                   |                   |                    | Pyth                     |                   |               |                   |                   | •                      |
| 5801 | CACCAACTTO          | OCCUATOGACE       | OCACROTITIO       | OFFITTION          | CCATCACCC                | <b>GCTCCTTGGC</b> | CGCGATGTTT    | ACCTOCACGE        | ATTCGCGCGC        | AACOCACCKA:            |
|      | OTCOTTOAAC          | COCTACCTEG        | CONCCCARAC        | CANARACAGE         | CTACCCGCG                | CCAGGAACCG        | GCGCTACAAA    | TCCACCTGCA        | TAAGCCCCC         | Tracoroccus            |
| 5901 | CATTCOODA           | AGACCOSTCOT       | <b>OCOCTCOTO</b>  | DOCACCACK          | CCACTICTICCA             | ACCIRCUOTING      | TGCAGGGTGA    | CANCOTCAAC        | actooracer        | Accherence             |
|      | GENAGOCOTE          | TCTGCCACCA        | COCCAGCAGC        | CCGTOGTCCA         | COTOTOCOCT               | TOTOCOCONC        | ACCITECTAL    | оптеслотто        | COACCACCOA        | TODYCYCCL              |
| 6001 | OTACOCOCTC          | GITTGGTCCAG       | CAGAGGCGGC        | COCCCTINAGO        | CCARCAGANT               | GENTALANGE        | GGITCTAGGTO   | contractor        | occoconcra        | COTCCACIANT            |
|      | CATCCGCGAG          | CANCCAGGIC        | OPENECUCE         | <b>GCGTGAAACTC</b> | <b>CETECHTETA</b>        | CHACKATICKE       | CCAGATCGAC    | <b>GCAGAGCAGG</b> | CCCCCANAC         | GCAGATACICA            |
| 6101 | AAAGACCCC           | OCCADENDO         | OCCCOTCORA        | GTACTATC           | THECATECTE               | GCAACTCTAG        | CRECTRICTEC   | CATOCOCOGO        | COCCANGCOC        | CCCCTCGTAT             |
|      | TTTCT0000C          | CCONCONCCO        | COCOCAGCTT        | CATCACATAG         | ANCOTAGGIA               | COTTICABATE       | GCTRIACGACG   | GTACOCOCCC        | OCCUPACIO         | COCCAOCATA             |
| 6201 | COUNTICACTO         | GOOGACCCCA        | TOCCATOCACO       | TOCOTTACCO         | CONTRACTOR               | CATCCCCCAA        | ATCTCOTANA    |                   | CICICIGAGE        | ATTCCAAGA"             |
|      | CCCAACTCAC          |                   | ACCOFACCC         | ACCEACTOCIC        | OCCITOCOCAT              | GTACGGCGTT        | TACAGCATT     | OCATCTCCCC        | DAGAGACTCA        | TARGOTTICTA            |
| 6301 | ATOTAGGGTA          | OCATICITICOA      | CCCCCANATIX       | TEXACECUCAC        | GTAATCHTAT               | ACTICOTOCG        | MASSAGECTAG ( | GAGGTCGGGA        | ccangottac        | TACOGOCOCY             |
|      | TACATCCCAT          | COTAGANOOT        | OCCCCTACG         | ACCOCCOCOTO        | CATTACK:ATA              | TCAMOCACGC        | recenedare.   | CTCCAGCCCT        | OCCITICANCO       | ATTRECENTED            |
| 6401 | CTOCTCTOCT          | COGRAGACTA        | TCTCCCTCAA        | GATCACCATGT        | GACTERADATE              | ATATOGETURE       | ACCICITIZANG  | ACCITICANCE       | TRACORCINST       | DAGACTTACI:            |
|      | GACGAGACGA          | <b>GCCTTCTGAT</b> | _                 | CTACCCTACA         | CTCAACCTAC               | TATACCAACC        | TOCCACCTTC '  | TOCAACTICG        | ACCOCAGACA        | CTCTCAMATO:            |

Figure 150

# PMRKAdSgng MER682

| 6501 | GCOTCACGCA               | COAAGGAGGC               | GTACKIAGTEG                                    | CACAGETYRAT                 | ACIACIC AGENT.  | CHICK PRINCE                  | PECACIFICAL<br>ACGRECAGAT  | CCCCCCRCAT                                  | GTCCAGGTT                       | TCCTTGATGA<br>AGGAACTACT    |
|------|--------------------------|--------------------------|--|-----------------------------|---|-------------------------------|----------------------------|---|---------------------------------|-----------------------------|
| 6601 | TOTCATACTT               | ATCCIUTCC                |  | ACAGCTCKICK<br>TISTICGAGCGC | CAACHCACHTA   | AAKTY-TTICT3C<br>TTITAK!AAGCS | GGTCTTTCA<br>CCAGAAAGGT    | GTACTCTTOG                                  | ATCCOMANCE<br>TAGCCTTTGG        | COTCOCCT**                  |
| 6701 | CCAACOCTAA               | CHCCCHAGCA               |  | CANCTECCO                   | #GCTAGGGGG<br>ACCATCCGTG                              | ALKTATICITE<br>TEXTAGGRAA     | TTCTACCACT<br>AAGATGCCCA   | AGCGCGTATO<br>TYCGCGCATAC                   | CCTCCCCCCCC                     | CTTCCG!AG                   |
| 6801 | CTCCACACCC               | TGAGGGGAAA               | CCACAGGGAC                                     | ACCATGACTT<br>TOGTACTGAA    | TGACKITACTES<br>ACTECATING                            | CATANACTIC                    | TCAGTGTCGT<br>ARTCACAGGA   | CCCATCCTICC                                 | CTGCTCCCAG                      | AGCAAAAAGT<br>TCGTTITTCA    |
| 6901 | CCCTCCCTT                | AAACCTTOCO               | GCTANACCGT                                     | CCCCCTTCCA                  | GACATCETT:  | AAGAGTATET                    | TTY CCCCCCCG               | AGGCATAAAG<br>TCCGTATTTC                    | TTGCGTGTGA<br>AACGCACACT        | TECHBANGS:<br>ACCCTTCCT     |
| 7001 | TCCCGGCACC<br>AGGGCCCTGG | TCGGAACGGT<br>ACCTTGCCA  | TOTTAATTAC<br>ACAATTAATG                       | CHOOOCOCC                   | AGCACGATCT<br>TCGTGCTAGA                              | ההדראאתהככ<br>הכאנידדדכפס     | CANCTACANC                 | TOCCCCACAA                                  | TOTARAGITIC<br>ACATTITICANG     | CANGAAGCG!<br>OFTCTTV:ff(1: |
| 7101 | CCCTACGGGA               | TONTOGNAGO<br>ACTACCTTCC | CANTITITAL                                     | ACTICCTCGT<br>TCAAGGAGCA    | AGGTGAGCTC<br>TCCACTCGAG                              | TTCACCOCAG<br>AAGTCCCCTC      | CIRRAGERAGE<br>GACITEGGGCA | CCACACTITC                                  | GGCCCAGTCT                      | GCAAGATHAAG<br>COTTCTACTK:  |
| 7201 | CCACCITICS               | CHOCHTACTC               | CTCCACAGGT                                     | CACTOOCCAT                  | TAGCATTTGC<br>ATCGTAAACG                              | ACCITCOTOCIC<br>TECACEAGEG    | GANAGMEET                  | ANACTOGOGA                                  | CCTATOCCCA                      | TTTTTCTGG<br>AAAAAGACC      |
| 7301 | CCACTACOTC               | _                        | CCCCAGANC                                      | Trecendend<br>Anggotese     | TCCCATCTAA  | CELANACGCCG                   | TACCAGAGCG                 | OCCIOCACITCA<br>COCCICTCAGT                 | CTAGAGGCTC                      | Atctccccca                  |
| 7401 | AACTICATOA               | CCACCATCAA               |  | TYCTTCCCAA                  | AGGCCCCCAVE<br>TCCCGGGGTA                             | CCAMETATAG<br>GCTTCATATC      | GICTCTACAT                 | CCTARGTGAC                                  | AAAGAGACGC<br>TTTCTCTGCG        | TCGGTGCGAG<br>AGCCACGCT     |
| 7501 | GATGCGAGCC               | GATCGGGAAG               | AACTOGATET                                     | CCCGCCACCA                  | ATTOCARTAG<br>TAACCTCCTC                              | Tractattea<br>Acceataact      | TUTCH CARANA A ACACCACTIT  | OTAGAAGTCC<br>CATCTTCAGG                    | CTCCCACCCC                      | CCGAACACTC                  |
| 1601 | GTOCTOGCTT               | TTOTAMANC                |  |                             | TCCACCACCT<br>ACCITCACCCCIA                           | GTACATCCTG                    | CACCAGCTTG                 | ACCTGACGAC COCOCACANO TOCACTGCTG GCGCGTCTTC | COCOCACANO<br>ACCOCACITIC<br>of | OAAGCAGAGT<br>CITCGICTCA    |
| 1101 | COCAATITOA               | GCCCTCGCC                | TCCCXCTTT                                      | CCGACCACCA                  | CTTCTACTTC  | CCGACCAACA                    | CCTTROACCGT                | CHARCTECTC GASGOGAGTT GACCGACGAG CTCCCCTCAA | GASGOGAGTT<br>CTCCCCTCAA        | ACCOPYCATY:<br>TRECACCTA!   |
| 7801 | CCTOOTOOTO               | COCCOCOCOA               |  | AGATCTCCCC<br>TCTACAGGCG    | מתמכנמת כמוד<br>כתכנע כסכיכת                          | CCTTCGAACT                    | TCACAACATC<br>ACTGTTGTAG   | GCGCAGATGG                                  | GAGCTOTCCA                      | TOGTICTICANA                |
| 1901 | CHCCCGCGGGC              | OPERBOTENO<br>CAGTECAGTE | OCOOMOCTC<br>COCCUTCGNO                        | CTSCACCITETE<br>GACGICCAAA  | ACCTCGCATA<br>TGCAGCGTAT                              | GACCAGTC                      | GCCCCGGCCGA<br>CCCCCCCGA   | AGATECAGGE (TETAGGTECA)                     | GATACCTAAT                      | Trecompage<br>Angireceed    |
| 8001 | TOGITOGROG               |                          | COSCUTCOAT GOCTTGCANG<br>GCCGCAGCTA CCGAACGTTC | AGGCCCCATC<br>TCCGGCGTAG    | ARGICUICATE CICCIOCHATOR<br>TECCOCCITAG (ROCITICATEO) | GACTACYATA<br>CTGATGCCAT      | מאכשהשבעם<br>מו אלשהששבע   | COCCOCCCC                                   | COCCOCATOTO                     | TCCFT(RGATY:<br>ARGANICETAL |

# PMRKAd5gag Mer682

| 8101 | ATOCATCTAA        | MCCCCTCAC         | <b>GCCATCCCANA</b> ; | いいいきがいい                             |                        |                    |                 |                    | CONCERNICACE       | GCGCGCGFF.    |
|------|-------------------|-------------------|----------------------|-------------------------------------|------------------------|--------------------|-----------------|--------------------|--------------------|---------------|
|      | TACGTAGATT        | TECHCACTO         | COCCURTATION         | האארון ארוקאראר<br>איזיין ארויקאראר | TECTTECTA              | GASCETOGGGG        | מאניוניוניוניני | ccerccccr          | CAGCCGC00          | CGCGGGCCGG    |
| 0201 | ASSAGETAGE        | CTCCCCCC          | TARCTICCTC           | הכיהאלתבינית                        | CTINTATION OF          | GTPGATCTCC         | TOWNSTANCE      | CCCTCTCCCT         | CANCACCACO         | ממכנונטאוניא  |
|      | TCCTCGACCA        | CONCOCOCOC        | ATCCANCIAC           | CGCTTCACCT                          | מישנים שלים            | CAACTAGAGG         | M TTAGACCG      | COGNUNCOCA         | cerciacnoc         | ברסמטככאריוי  |
| 8301 | <b>GCTTGANCCT</b> | CAAAGAGACT        | TEGACACAAT           | CAATITICAGE                         | CHESTITICAGE           | فالازميان بالمادان | CIC MANATUTE    | CTYCACOTOT         | CCTCAGTTGT         | CTTCMTAGG"    |
|      | CGAACTTOGA        | CHITCICICA        | AGCTGTCTTA           |                                     | OTTAMOLICA CAGINACINIC | CHICHERECES        | CCTTTTAGAG      | GACOTOCAGA         | GGACTCAACA         | GAACTATCC:    |
|      |                   |                   |                      | E                                   | 111                    | ,                  |                 |                    |                    |               |
| 8401 | GATCTCOOCC        | ATGAACTOCT        | CONTETETE            | _                                   | CTCCTCGAGA TCTCCGCGTC  | נישיניונימכוב      | CACOCHICCO      | GCGAGGTCGT         | TOGRANTOCO         | CCCCATGAC     |
|      | CTAGAGCCOG        | TACTTGACGA        | <b>OCTAGAGAA</b>     | _                                   | NUNCTURE               | CCCCACCCAC         | CTYCCACCGC      | COCTCCAGCA         | ACCTITACGE         | CCCATANCTIC": |
| 8501 | TOCGNONGO         | COTTORAGGCC       | TCCCTCGTTC           | CAGACGCAGC                          | TETAGACCAC             | GUCCECTICE         | GCATCGCCAG      | COCCICATIOAC       | CACCTOCOCO         | AGATTOAGE "   |
|      | ACOUNTING         | GCAACTCCGG        | AGGGAGCAAG           | grenococch                          | ACATICITISCITIS        | CHRICHRIANIC       | CCTAGCOCKIC     | <b>ACCIONATION</b> | UTGGACGCGC         | TCTMCTCO.     |
| 8601 | CCACOPUCCO        | OCCOANGACO        | <b>OCOTAUTITE</b>    | OCACOTECTO                          | AAAGAGGTAG             | TITCACACATIVES     | nacceaning      | TTCTGCCACG         | ANGRAGTACA         | TAMCCCARC:    |
|      | CONCEACOCC        | CCCCTTCTGC        | CCCATCAAAG           | CONCCOCOAC                          | TTTCTCCATC             | MCTCCCACC          | ACCECCACAC      | AAGACGGTGC         | TECHTCATOT         | ATTOOOTCCH:   |
|      |                   | <b>≖</b> {        | Ecoffy               |                                     |                        |                    |                 |                    |                    |               |
| 8701 | TCCCMCGTG         | DATTECOTTEA       | TATCCCCAA            | OCCUTCANOD                          | CCCTCCATCA             | הכדורהדאמא         | GICCACOGCO      | ANCIFICADAA        | ACTOSCAGIT         | GCCCCCCAC     |
|      | ACCOPTICCAC       | CTANGCAACT        | ATAGGGGGTT           | CCCCAGNITCC                         | <b>GCGAGATACC</b>      | GRAGGATCTT         | CAGGIGCCGC      | TICAMCTITI         | TOACCCTCAA         | כככסטספבונו   |
| 8801 | ACCONTANCT        | CCTCCTCCAG        | AAGACGGATG           | MOCTOGREGIA                         | CAGTOTCOCO             | CACCTURING         | TCAMAGGCTA      | CAGGGGGCCTC        | TICHCFICT          | TCAATCTCE:    |
|      | TOCCAATION        | COMOCHOCIC        | THETECCTAC           | TCGAGCCGCT                          | GTCACAGCGC             | GTGGAGCGCG         | AGTITICCGAT     | GTCCCCCGGAG        | AAGAAGAAGA         | ACTTAGAGGA    |
|      |                   |                   |                      |                                     |                        |                    |                 |                    | Spil               | 2             |
| 8901 | CFFCCATAAG        | <b>GOCCTCCCCT</b> | refrement            | CTCCCCCCCC                          | TYXCACKCK              | GCGACACAGC         | GGCGACGACG      | <b>GCOCACCOGO</b>  | Абосоотса          | CAMBEGETE     |
|      | DAAGGTATTC        | CCCCACCCCA        | ACANGANGAN           | פעיבפכבפכב                          | Anconcretors           | CCCTOTOCCG         | COGETRACTICS    | COCOTOOCCC         | TCCGCCAGCT         | OFFICATIONS   |
| 9001 | GATCATCTCC        | CCCCCCCCAC        | OCCCATION            | CTCGGTTACG                          | GEGEGGGGGG             | Transpoora         | GCACACATICO     | AAGACGCCGC         | CCONCANGIC         | CCCATTAICE    |
|      | CTAOTAGAGG        | OCCCCCCTG         | CCCCCTACCA           | GAGCCACTGC                          | CCCCGCCCACCA           | ACAGACACCC         | CCCCTCAACC      | Tretocoded         | OCCUBTACAD         | OCCCAATACK:   |
| 9101 | GETTOOCCOOC       | OCCIOCCATO        | COCCAGGGAT           | ACCOCCTAA                           | CGATTICATET            | CAACAATTAT         | TUTOTAGGTA      | בובנמכנפכ          | GACCCACCTO         | ACCCANTERY    |
|      | CANCEGETE         | CCCACCOTAC        | <b>accorcceta</b>    | TOCCGCGATT                          | GCTACGTAGA             | GTTGTTAACA         | ACACATOCAT      | омассавсав         | CICCCTOCAC         | TCCCTCACC     |
|      |                   |                   | XOrd                 |                                     |                        |                    |                 |                    |                    |               |
| 9201 | CATCOACCOG        | ATCODAMAC         | CTCTCGAMA            | ACCIONTETAA                         | CCAGTCACAG             | TCCCAACCTA         |                 | COTOBCOORC         | <b>OCCADE GOOD</b> | OCCCCITCCIA:  |
|      | <b>GTACCTOCCC</b> | TAGGCTTTTG        | GAGAGCTCTT           | TCCGCAGATT                          | GGTCAGTGTC             | ACCCTACCAT         | þ               | GCACCGCCCG         | CCCICCCCC          | CCGCCAGCCC    |
|      |                   |                   |                      |                                     |                        |                    |                 |                    |                    |               |
| 9301 | OFFICIAL          | acadagatac        | TOCTGATCAT           | GTANTTANG                           | TACKACIANT             | TYTAGACYBCCG       |                 | AGAAGCACCA         | TITLE CT TOO       | Tecanectice   |
|      | CAACAAAGAC        | COCCTCCACO        | ACGACTACTA           | CATTANTHE                           | ATCUTACOGA             | ACTOTOCOGO         | CTACCARCTO      | TCTTCGTGGT         | ACAGGAACCC         | ACCCCCCACC    |
| 9401 | TUAATUCOCA        | occoorcoc         | CATTRCCCCAB          | -                                   | GACATICITION           | CACASTICATIVE      | TAGTAGECTT      | CCATGACCT          | TICTACCOOC         | ACTICITIET    |
|      | ACTTACOCOT        | CCCCCACCC         | CTACCOCCTC           | CGANGCAAAA                          | CTCTANCCOC             | CHCCAGAAAC         | ATCATCACAA      | COTACYCOGA         | AMGATGGCCG         | TGANGRAGIN    |
| 9501 | CRECTICE          | TTOTOCHOCA        | TCTCTTCCAT           | CTATRICATION                        | تعصينون                | CACTTRACE          | CLYVICTORIES    | ccerement          | CCCATGCOTO         | TOACCCCOAN    |
|      | CACCAACCAC        | AACAGGACGT        | AGAGAACGTA           | GATAKA:GACG                         | כנואבנואכנוטב          | CTCAAACITED        | CATCCACCC       | CCCACACAACCA       | COCTACOCAC         | ACTOGGIACT    |
| 9601 | <b>ACCECTEME</b>  | GOCTGAAGCA        | CAXICTACOTC          | CHATCHACTACO                        | CCCTATCACTA            |                    | CHICACCTOC      | GTGAGARTTAG        | ACTURINAGIC        | APCCATIFICE   |
|      | COCCAGTAC         | CCGACTTCCT        | CCCCATCCAG           | ccacrorrac                          | CCCACCCCAT             | TATACCOGAC         | CACCTCCACC      | CACIFCCCATC        | TGACCTTCAG         | TACGENCACO    |

Figure 15F

| STATE:<br>SATEN:<br>PAGAC:   | ארכדונ   | 3000C                        | بورو <b>د.</b><br>بورود:                       | JACCT!                          | מאנים<br>מאנים                                       | MG#:                           | TOWN 1                   | LEGG CO                   | וניארני<br>וכידיאר       | ICACA:                    | ACCTC                     | MCCO                      | MCAA                     | וניו יניו<br>זכיזיפיני<br>זכיזיפיני                                      |
|--|--|------------------------------|--|---------------------------------|--|--------------------------------|--------------------------|---------------------------|--------------------------|---------------------------|---------------------------|---------------------------|--------------------------|--|
| TCGGTGTAFFG  AGCCACATURE   |  | r acconcinica<br>A conceence |  |                                 |  | A GCATTANGTHE<br>F COTAATTCAIT | CCCCCANGE                |                           |                          |                           |                           | C CCCCCAACCO              | CCTTTAACAA               | TAXCAMENTES  |
| CTGCCIACAGC GACGCTCTCG   | -  | TCCAGGTGAT<br>AGGTCCACTA     |  |                                 | CACAACOOGO   | GAMAGCGANA                     | TOCOGCOING<br>ACCCCCCTTO | COGTOCTOCO                | CCCCTOTAGO               | CCCCTACCAC                | AGGGAGAGAGA<br>TCCCTCTCCT | TOAGCCCGA<br>ACTCGGGCT    |                          | TTTTGGGTTT<br>GCCCGAGGC  |
| CCACTGGGC  | GTTTTCACG                                      | TACCTOGACA                   | TUGGGACGCT                                     |                                 | TOCOACOTOA   | GOTTAGGCTG                     | CCTGCCGGAC               | CAGATGCÁTC                | CCTCACCACC               | CCCCGGACCGC               | COCCOMOCTOCO              | AGGAGGACTT<br>TCCTCCTGAA  |                          | GCGCTOGAGC<br>CCCCACCTCG<br>ACATAGTAGA                                   |
| NACHRIACTAG TTAACKGICT<br>TTGCCTOSTC ANTIGCCAGA  | TREPLEATING CONTACTOR  FORTY                   | CCCCTACTAT ACCUTACATOR       | TGCTCCATER                                     | CTCCTCGATA                      | ANY CCAPATED THE THE THE THE THE THE THE THE THE THE | CMATCHANGE                     | CACAGAGCCT               | TTGCTTTTCC                | CCTCCTACCG               | AGGAGGGCGA<br>TYCTYCCCGCT | GAACCTGTTT                | THECTGCGG<br>AACGACGCGC   |                          | ATCTFTTCTCA CTTTTTANGE<br>TAGACACCCT GAAACATTCO<br>ATTCACCTAT GCFTTCCTAA |
|  |  |                              | CCCCTTTTTC                                     | CTPLCGINGT                      | CCCCTTTTTC   | Traccorera<br>Accoscocec       | CCCCCGGTTC               | AGCCCCTTTT<br>TCGCGGANAA  | ACCETCCC!!               | CTREACTTRE<br>GACCTRAAACC | TCCCCCCCCTA<br>ACCCCCCCCT | Tr:GCGAGCCG<br>AGCGCTCGCC |                          |  |
| AGTTCGCCAT   | CCATCARCAT CATACHCO CO                         | GOSCICTAGA AGGITGIATA        | CKATTECKAN TETTREGERA<br>GECANGETET AGANGGEOTE | AGCONGINET<br>TEGECECTICA       | CONTRACTOR   | TTTTKTCCAC<br>ANACCOTTO        | ACTUCACOCCCT             | AACAGGGAGG<br>TTYSTCCCTGC | CATCACACAC               | CCCCCCACTAC               | CACCCCATCC                | ATOGCCTGAA<br>TAGGGGACTT  |                          | GGACTIONTING<br>CCTGACTINIO<br>ACAMPINING                                |
| OPERATORIAN CONTRACTORIAN CONT |  |                              |  | AGAGCCTR:TA<br>TCTCGGACAT       | GTCTACCATC   | GCGCTAGCTT                     | CCAAGGGTTG               | TTCCTCCGGA<br>AAGGAGGCCT  | CAGCGGCAGA               | CCACAGCCCA                | TOATACGCGT<br>ACTATOCGCA  | GAGCTOCOAC                | CCCCCCCACCT              | GOTOGCTATA<br>CCACCGATAT<br>CACAGCATAT                                   |
| OGTATOCOCC CONSTIGATO (TICTAAGINE) CCATACOCOG OCACAACTAC (INCATTERA) MANAGEM   | GTANGCCCTC GAGTCAAATA<br>CATTCGGGAG CTCAGTFTAT | CCCCAGGCCC                   | OTCOCKANGO<br>CAGCGCCTGC                       | <b>CACOTITICS</b>               | GCCGTCCTCC<br>CGCCAGGCCA                             | CCCCCACCA                      | CCCANTRAGA               | CCCTTCCAAA                | AGACCAAGAG<br>TCTCCTTCTC | GAACCCCCCC                | AGCTHANGCG<br>TYCGACTTCGC | COCAROGOGO                | CACGTGGCRG               | CCCCCCCACGA<br>CCCCCCCCC<br>TATAGTTCCAG                                  |
| GGTATGCGCC C<br>CCATACGCGG G   | GATTCGGGAG                                     | ARGOTOGCCO<br>TCCCACCOCC     | CGCGCGCGAAA                                    | <b>GCTCTAGAGG</b><br>CGAGATCTGG | CCCOTATCCO   | Trechogede<br>Anotecede        | ACATEGECCT               | TOCANGACCC                | AGCAGCGGCA               | TOOTOATTAC<br>ACCACTAATO  | CCANGGOTGC                | GAAAGTTCCA                | COCOCOCOCO               | ACCCTIONOC<br>TOCCOACACC<br>ACCTUITICCT                                  |
|  | TOAGNCGCOA<br>ACTCTGCGCT                       | CCCOOTCOCA                   | GACCACCTCC                                     | AATCOTTOAC<br>TTAGCAACTO        | CCCAAGCTCG   | MACCOMOD                       | COACCOACC                | CICCCCOTCA                | CCCCCTCCTC               | CCCCACCACA                | TEAGCOGCAC                | ATCCCCCARG<br>TACCCCCTRG  | GUATTAGTCC<br>CCTAATCAGG | CCACOTOCOF<br>OOTGCACOCA<br>CTCATOOCOC                                   |
|  | 1086   | 9901                         | 10001  | 10101                           | 10201  | 10301                          | 10401                    | 10501                     | 10901                    | 10701                     | 10801                     | 10601                     | 110011                   | 11101  |

|   |                 | and the same          | -depresentation |                        |                     |                      |                     |                        |                    | Sept. B. Co. Co. Co. Co. Co. Co. Co. Co. Co. Co |  |
|---|-----------------|-----------------------|-----------------|------------------------|---------------------|----------------------|---------------------|------------------------|--------------------|---|--|
| 11301                                   | TCCAPTICAL      | •                     |                 | かいかいかいいり               | CHICK WAY THE       |                      |                     |                        | TATTICCATIC        | AATCISIACIT                                     |  |
|   | AGCTANACTA      | THOTAGOAC             | GTCTCCTATC      | ACITACITACIT           | נומוז: זונונון      |                      |                     |                        | 2000000            |   |  |
| 11401                                   | CAAGITTAC       | CCCCCCAAGA            | TATACCATAC      | CCCTTACCTT             | CULINIMIACA         | ACKTACIONANA         | いとしていること            | _                      | CCATCCCCCT         | CAALSTIN  |  |
|   | CITICAAANO      |                       |                 | COCHATOCAA             | CATATATATATA        | TYTYCATT             | CTANACTICINIC       | ANGATOTACO             | COTACCOCOA         | כבוניבעכני י                                    |  |
| 11501                                   | POLLENS AND DE  |                       | COTITATOR       | AACGACTCCA             | TEXTA MANGE         | CYTTCANTCTTC         | MICCORPINGE         | DUCARCTICAD            | CONCCOCUAG         | CIVINITIZACA                                    |  |
| 10011                                   | Percia McTerrat |                       | GCAAATAAT       | TRETEGEST              | ACKSTICT TOUGH      | <b>GCALTTORICAC</b>  | Trynoceaeco         | COCTCGAOTC             | <b>GCTOGCCRCTC</b> | GACTACOTATE                                     |  |
| •                                       |                 |                       | CONTRACTOR      | CONTRACTAR             | אניאכנינטאני        | TCCTACTTR            | Accessorac          | TEACCTECCE             | TOGGCCCCCAA        | <b>OCCUPACIOCO</b>                              |  |
| 1001                                    | COCACOTTIC      |                       |                 | CRCCKRCTATC            | TUTCOGCTC           | ACCIATICANAC         | TOCOCCCACG          | ACTGGACGCG             | ACCCCCCCCTT        | COSCHOCOC:                                      |  |
| 11701                                   | CONTRACACTOR    | _                     | GACCTGGGCT      | CHCCGGTCGCA            | تتنديريون           | CTGGCAACGT           | COCCACCATO          | GARCAATATO             | ACGAGGACGA         | TOACTACGAG                                      |  |
|   | CONCETCOO       |                       |                 | CCGCCACCGT             | ממאנענענענענ        | GACCGTTGCA           | OFCGCCGCAC          | CTCCTTATAC             | TOCTOCTOCT         | ACTCATOCTC                                      |  |
|   |                 |                       |                 |                        |                     |                      |                     |                        |                    |   |  |
| 10811                                   | CERCACORCE      | a archatheth          | AGCGGTGATG      | TITCEGATEA             | CATCATGCAA          | GATHCAACTG           | ACCERGOOOF          | 000000000              | CTGCAGAGCC         | AGCCGTCCCP (                                    |  |
|   | OGICICAGO       |                       |                 |                        | CTACTACGIT          | CHACITHECE           | <b>TRIGRECTANCA</b> | COCCCOCCOCC            | GACGICTCOG         | TCGGCAGGC                                       |  |
| 11901                                   | COTTANCTIC      |                       | _               | CATOGACCIAC            | ATTATETORS          | PRACTYSCALG          | CAATCCTMAC          | OCUPTOCOOC             | AGCAGCCGCA         | GCCCNCC#1                                       |  |
|   | CCAATTOROC      | _                     | -               | GTACCTGGCG             | TAGTACAGE           | ACTIGACIGICISC       | GTTAGGM:10          | COCANGOCCO             | regregator         | CCGCTTCCC                                       |  |
|   |                 |                       |                 |                        |                     |                      | - E-                |                        |                    |   |  |
| • 000                                   |                 | A STATE OF THE PARTY. | - astrony       | Carrier Carrier AA     | ACCCCACACA          | CGAGAGGTG            | CTGGCGATY:G         | TAMACCICCT             | GGCCGAAAAC         | AGGCCATCT                                       |  |
| 10021                                   |                 |                       |                 | -                      | TOCOCTOCOT          |                      | CACCCCTACC          | ATTTOCOCCA             | ccacering          | TCCCOOTAG                                       |  |
| • |                 |                       |                 |                        | acontrop.           |                      | GCGCCAACGT          | GCNGACCAAC             | CTOSACCOSC         | TROTROGGGA                                      |  |
| 10171                                   | GALLEMED        |                       |                 |                        | COLOCACCGA          | CCANTICITIVIT        | CACCATTACCA         | ceremonro              | GACCTGGCCG         | ACCACCCCC"                                      |  |
|   | CCARACTAC       | _                     |                 |                        | Canada Property     |                      | COPTICACION         | AACTEC                 | TUAGTACACA         | OCCCOCCAN'                                      |  |
| 12201                                   | TOTOCOCOA       | _                     | -               |                        | ביא(פוסגיור אישו. ר | ACCOUNT OF           | CCAACCERCIAT        | THECOMAGG              | ACTENTICION        | COCCIONT  |  |
|   | ACACCOCCTC      | C COOCACCOCO          | TCCCACTCGC      | ACCEPTION              | OPECCE: 100         | אררריישראו           | רכוסורסוניייו       |                        |                    |   |  |
| 12301                                   | OTOCCOCOO       | O GACAGGAGGA          | CTACACCANC      | PTTOTOTOGCO            | CACTGCGGCT          | AATGGGGGG            | מאמאכאכנפכ          | AANGTGAGGT             | GINCCAGICT         | GAN.I. MAN.                                     |  |
| •                                       | CACOGCCCCC      | c croncerect          | _               | GATGTGGTTG AAACACTCGC  | GIRCACCCCTA         | TTACCACTGA           | CICIOTOCCO          | TTTCACTCCA             | CATGGTCMA          | ררואורואו                                       |  |
|   |                 |                       | Psd             | _                      |                     |                      |                     |                        | •                  |   |  |
| 12401                                   | ATTENDED        | A GACCAGIAGA          |                 | CAMPRICATIC NANCOGTNAA | CCTGAGCCAG          |                      | ACTROCAGES          | cctctccccc             | מתוכנים            |   |  |
| 1                                       | TAAAAAAGG       |                       |                 | TUTCHENT               | GGACTECATIC         | CCANAGETETE          | TOMOGRACIC          | CCIACACCCCC            | CACOCCCCAAG        | Greence   |  |
| 12501                                   |                 |                       |                 | CAACTORIC              | CTGFTXTTCC          | TOCTAATAGE           | GC TOTAL ACC        | CACACTOCCA             | acoratecea         | GGACACATAC                                      |  |
| 10001                                   |                 | . ~                   |                 | GTTCAGCIACG            | מאניאאכמאכש         | ACCAPTANCE           | CCRRIMACTIVE        | CHCTCACCGT             | CCCACAGGGGC        | CCROIGIAN                                       |  |
| ,                                       |                 |                       |                 |                        | AGGINGLATICE        | CATALTEAUSTAT        | NUTTICCARG          | ACATTACAAO             | TOTCAGCCGC         | מכנוכרותפתולכ                                   |  |
| 12601                                   | CTAGGICAC       | T TOTAL MERCAL        |                 |                        |                     | CCTAXTECUTA          | THANAGETEE          | TCTAATGTTC             | ACAGTEGGGG         | COCONCCCO                                       |  |
|   | CATCLAGIC       | ě                     |                 |                        |                     |                      |                     | •                      | Pireple            |   |  |
|   |                 |                       | CROST BROTT     | TANALTACE              | CALTENACYANG        | CTARCTARCAGA         | AGAITCCCCTC         | OFFICEACAGT FFINACAGES | <b>FFNANCAGED</b>  | ACCARICAGES                                     |  |
| 12701                                   | MODELLACAL      |                       |                 |                        | CCACTINITIES        |                      | TUTAKKARAG          | CAACGTGTCA             |                    | PCCACCTA(%)                                     |  |
|   | 1000            |                       |                 |                        | STATERA             | CONTRACTOR OF STREET | CARCITICANO         | CTRICACATICA           | הכמכסכתה           | CATCSTAACTT                                     |  |
| 12801                                   | GATTITIGEGE     | O ATGCACGTCG          | TCTCTCCACTC     |                        |                     |                      | מוכנאניאנינאנ       | GACKTIVITACT           | accesect           | GTACCTTGGC                                      |  |

Figure 15H

# PMRKAd5gag MERGR2

| 12901 | OCCATOTATO       | CETENANCER                              | OCCUPITATE         | AACCERCICAA               | TYSTACTA           | מיאראיאנא                 | Christophica        | ACCECCAMOTA            | TPTCACCAAT        | GCTATCTTCA        |
|-------|------------------|---|--------------------|---------------------------|--------------------|---------------------------|---------------------|------------------------|-------------------|-------------------|
|       | CCOTACATAC       | GOAGTITOGC                              | COCCAMITAG         |                           |                    |                           |                     | TYPECACTCAT            |                   |                   |
| 13001 | ACCCRCACTO       | GCTACCUCCC                              | cetentrici         | NCACCURRAN                | ATTRICACACETES     | אדירניאנידיני ניייהאנידאר | ACGATGRATT          | CENTERODORE            | GACATAGACG        | ACACACOTOTA       |
|       | TOCCOTOAC        | : CGATGGCGGG                            | GOACCAAAAA         | ו אאווטכנינכ              |                    | CXXICTECAT                | TKICTACCTAA         | GGAGACCCTG             | CTOTATCTGC        |                   |
| ,     |                  |   |                    |                           |                    |                           |                     | HingRil                | •                 |                   |
| 13101 |                  | CCGCAGACCC                              | TOCTACACIT         | . פנישעכשינוני            |                    | GAGCAGGCAG ACCACACACT     |                     | GCGAAAGGAA AGCTTCCCCCA | GGCCNAGCAG        | CPTGTCCGAT        |
|       | AAGGGGCOFF       | r occoncrado                            | ACCUATOTICAA       | בנידוסוכככם               | : ديدنويد.د.يد     | דיניניכנינינות            |                     | TCGAAGGCGT             |                   |                   |
|       |                  |   | ٠                  |                           | Ulineliti          |                           |                     |                        |                   |                   |
| 13201 | CTAGGCCCTC       | COCCCCCCC                               | <b>OTCANATICET</b> | . AGTAGCCCAT              | TTCCANCITT         | CATACACATACT              | CTTACCAGCA          | CTCCCACCAC             | CCCCCCCCCCC       | CTCCTOTICO:       |
|       | DATICCOCCAC      | 300000000000000000000000000000000000000 | CAGTCTACGA         | TCATCOCCTA                | <b>AAGCTTCC:NA</b> | CTATCCCAGA                | GAATGGTCGT          | GACCOTOCTO             |                   |                   |
|       |                  |   | Pst                | -                         |                    |                           |                     |                        |                   |                   |
| 13301 | ACCACCACTA       | I OCTANACANC                            |                    | לכטכיוזיני מאכה אכת אהרים | CGNANANAC          | CHECCHECES                | CATTROCCA           | CAACGGGATA             | GAGAGCCTAG        | TOCACANIA1        |
|       | Techecteat       | · DOATHOTTO                             | AGCGACGACG         | TOGGGGTCCC                | GCTTTTTTG          |                           | GTAMORETT           | GITGCCCTAT             |                   |                   |
| 13401 | OAGTADATOG       | AMGACOTACO                              | COCAGGAGGA         | CACCOLOTTO                | CCMPICTIGG         | GUCCIACCAC                | COSTECUTEAN         | AGGCACGACC             | OTCARCOORD        | reromana          |
|       | CTCATCTACC       | : TICTOCATUC                            | <b>OCOTOCTOR</b>   | GITCCCTGCAC               | פטערינינאטפפ       | כנאנשטנאנפונם             | GOCAGCACTT          | тесотостов             | CAGTCOCCC         | AGACCACACC        |
| 13501 | GAGGACGATG       | ACTUGGCAGA                              | COACARCARC         | GICCICCACT                | TRECACTANG         | TRICARCCCG                | TPROCACC            | TTCGCCCCAG             | CCTCCCCCACACA     | ATCTITTAM         |
|       | CTCCTGCTAC       | : TOAGCCOTCT                            | <b>GCTOTICGREG</b> | CAGGACCTAA                | <b>ACCUTECATE</b>  | ACCGFTRAGG                | <b>NANCYSCETTIO</b> | AAGCGGGGTC             | COACCCCACT        | TACAMATIT         |
| 13601 | ANANAANAAA       | GCATGATGCA                              | ANATANNAN          | CTCACCAACA                | CCATTACACC         | GACCUTTRICAT              | TETCTTGFAF          | TCCCCTTART             | ATCCCCCCCC        | COCCIONTOTA       |
|       | erterrents.      | COTACTACGT                              | TTTATTITE          | GAGTRIGHTCC               | COSTACTION         | CINCOLACCA                | ANAGANCATA          | AGGGGAATCA             | TACGCCGCGC        | <b>OCCOUTNOAT</b> |
| 13701 | TOAGGAAGGT       | · concener                              | CCTACCIACAG        | TOTOSCICACE               | <b>GCGGCCCCAG</b>  | TRACEGUAGE                | CKTROOPTICE         | CCCTTCGATO             | CTCCCCTOOA        | CCCCCCTTT         |
|       | ACTCCTTCCA       | GOAGGAGGGA                              | GGATGCTCTC         | ACACCACTCG                | COCCOCOGIC         | ACCCCCCCCG                | CGACCCAAGA          | GGGAAGCTAC             | <b>GAGGGGACCT</b> | GOCCCCAA          |
|       |                  | Kind                                    |                    |                           |                    |                           |                     |                        |                   |                   |
| 13801 | OPOCCACCOC       | GGTACCTOCO                              | <b>GCCTACCORD</b>  | GRIGAGAAACA               | <b>ACATICOUTTA</b> | CTCTGAGTTG                | GCACCCCTAT          | TOBACACCAC             | CCGTGTGTAC        | CTGGTGGACA        |
|       | CACCOACCC        | CCATGGACGC                              | COGNICCICC         | CCCTCTFIGE                | CCTAGGCAAT         | CARIACTERAC               | CCTCCCGATA          | ACCTURGED              | COCACACATO        | GACCACCTOT        |
| 13901 | ACAMOTICANO      | _                                       | TCCCTMARCT         | ACCAGAAACGA               | CCACAGGAAC         | TTICTICACCA               | COUPLATIE           | AAACAATGAC             | TACAGCCCOG        | GGGAGCCAAG        |
|       | TOTTCAOTIC       | CCTACACCGT                              | ACCCACTTOA         | TROTETTRICT               | <b>GOTOTOOTTO</b>  | AAAAACTKKTT               | CCCAGTAAGT          | TINCHINCIO             | <b>ATGTCGGGCC</b> | CCCTCCCTP.        |
| 14001 | CACACACACC       | ATCANTCTIO                              | ACGACCOGTC         | GCACTCARARC               | CKICCACCTRIA       | AAACCATCCT                | GCATACCAAC          | ATGCCAAATO             | TOANCOAGTT        | CANTITITACC       |
|       | GIOTOTOTO        | TAGTTAGAAC                              | TGCTCGCCAG         | CONTRACTOR                | CCRITICALL         | TTTCCTACTA                | CGTATGGTTG          | TACOGETTEAC            | ACTITOCTICAA      | GTACAVATA         |
| 14101 | ANTANGITIA       | AGGCCCCCCCT                             | GATOCTOTCO         | COCTROCCTA                | CTAAGGACAA         | Transmissan               | CTCAAATACG          | AGTGGGTGNA             | OFFICACOCTO       | CCCCANATACA       |
|       | TTATTCAAAT       | TCCGCGCCCA                              | CTACCACAGC         | OCCUNICACY                | GATTCCTVITT        | AGTECALITIC               | GACTITATOC          | TCACCCACCT             | CAAGTGCGAC        | GGCTCCCGT         |
|       | -                |   |                    | •                         | Pverif             |                           |                     | •                      |                   |                   |
| 14201 | ACTACTOCOA       | GACCATOACC                              | ATMGACCTTA         | TOANCANCAC                | GATCHINGAG         | CACTACTEGA                | AACTGROCAG          | ACAGNACOCO             | GTICTOGAAA        | RECONCATEOR       |
|       | TOATGAGGET       | CTOOTACTOO                              | TATCTICAL          | ACTIVITIES                | CTACH:ACC:TC       | CHEATGAACT                | THEACCOCHE          | TOTOTAGE               | CAAGACCTTT        | CIXTITABEC        |
| 14301 | <b>OCTANGETT</b> | GACACCCCCA                              | ACTTCAGACT         | CONTITION                 | ددىسىسى            | THE THENCH                | CCCTTAXXXTA         | TATACAAACO             | MOCCITICGA        | TCC NGACATU       |
|       | CCATTTCAAA       | CTGTGGGCGT                              | TOMOTOTOR          | נתיושעריםם                | שהיויאיאצא         | נאטאטאטאט                 | COMMUNICATE         | ATARCHING              | TTCOUNGET         | MAGICINETAN       |
| 14401 | ATTTOCTOC        | CACCOATGCGG                             | CONTRACTIC         | ACCCACACRCC               | CICTERACK:AA       | CHRISTIAN                 | ATCITICANCE (       | COCARCCOTT             | CCAGGAGGGC        | TTTAGGATCA        |
|       | TAMACCACC        | <b>STECTACOCC</b>                       | CCACCTRIANG        | TOTALIST                  | COCACTOCAT         | מאכאכננט                  | TAGRECOTTES         | CCCTTCCCAA             | <b>GGTCCTCCCG</b> | みんれたですからず         |

Figure 15I

| CCCCACTOR A POCCATACTOR ACCOLAGGE GAGAAGCE CTCCATACTOR CATACTOR CATACTOR ACCOLAGGE ACC | GCAG' FG/FT'A<br>CGTC(ACCAT  | GATGACCA(C<br>CACTCCAAGA<br>GTGAGGTTCT<br>Asd   | AAACCUC();                | AGTCCAGCON<br>TCAGCOTCÓCT<br>TTTTGAGCAA      | AAAACHCHT<br>ACCAACACC                 | CCACCACT ANATOMEN TITTACTICT                           | GCACCTA M.C.I.               | CCCCATCCATT<br>GCCCCCATAA<br>CCCCAACTAAA | ACCITIVATOR ANGAGATORY TICTORATORY                           |
|--|------------------------------|---|---------------------------|--|--|--|------------------------------|--|--|
| CHACAGGGCG CO<br>CTAGACCGCC CO<br>TRAACGATCA TA<br>ACTTGCTAGT A<br>ACCCGAGGTC G  | ACCCAGINCE G<br>TOGGICATOG C |   | GAGAACCAGA T              | CCATCGGAGG A<br>CGTAGCCTCC T<br>GAGCCGCACT T |  | _  | CECECACOTE G<br>GEGEOTECAG E | CCCCCCCCCCC G                            | _  |
| AGATGACACC<br>TCTACTGTGG<br>GTCGACGTCCTGT<br>CCCCCTGGT<br>CCCCCTGGCACGCGA  |                              | TOTAL COLOR OF TOTAL | TCGCTTTCCC<br>AGCGAAAGGG  | CTOCOCAACA<br>GACGCOFTGT<br>GCOTCCTATC       |  |  | CCTOCCTAAC<br>GCACCAATTO     | CCACCCCGGC<br>GCTCCCCGGC                 |  |
| CONGCTTONA<br>GCTCGAACTT<br>ANTGTAKGCG<br>TTACGTAYAG<br>GCTAGGCGA  |                              | TCCTOACTTACAT ARBACTTACAT CCCGTTACATAG GRCCACCACC   | ACTROTICAN<br>TOCACAGGIT  | GACGCTANCG<br>CTYCGATODC<br>GTCTCGCCGC       |  | GCCCACCACC<br>CCCCTCGTCG<br>ACCCTCGTTCC<br>TCTCACCACG  | محدمحدم<br>محدمحدم           | OTCCACACCA<br>CAGGTCCCCT                 |  |
| GTTPACCARI<br>CXSATVSTVC<br>CACVINCENC<br>CRYSHITSIA<br>MYYPHTTHA<br>TXTHCIARCTT   |                              | TECTTRICAC<br>ACEANACGEG<br>CAECAACTET<br>GEGETIGAAA  | TCTCTGACCC<br>AGAGACTGGG  | CAGATCACGG<br>GTCTAGTGCC<br>CCTGGGGCAFA      |  | COCCATACATE COCCATICAG CCCCATICAG CCCCATAGTC           | CREGATOCOC                   | ACTAROGRATIC                             |  |
| CRIANGINGAC<br>CCTACACTITO<br>TYCCAACTITA<br>AGGITGAACTI<br>AGGITGAAACTITA   |                              | TEATGRANCE<br>AGTACCTTAGG<br>CGCCGCTCTA   | CCAGTITACC<br>GGTCAAATGG  | CCFFCTCTCA<br>GCACGAGAGT<br>TTFACAAGGC       | _                                      | CACAAACGCG<br>GAGAAGACGC<br>CACAACCAGGG                | CHACACTACC<br>CACCGTGACGG    |  | TECTACIONE TO<br>ACCESTANO<br>TECANOCICACIO<br>ACCETACICACIO |
| CCCCACTESTS QCCCTTACAA GRIANTAGAAC CCTTCCTTTG ANGCCGCAC  |                              | CCCTACCC<br>OCCTTACCC<br>TTCCCCTCCA<br>AACGCCTACCA  | AACTCATCCG<br>TTGAGTTAGGC | TGAAAACOTT<br>ACTTTTOCAA                     |  | CTRACOCCOCC<br>GACCCCOCCAC<br>CCAGTGTTCA<br>GGTCACAGGT | COCCOCCOACC                  |  | GTGTATARAGE<br>CACATARCE<br>GTGTATCTA                        |
| CCATTACATTC<br>CCATTATAGG<br>GCANTANGGG<br>CUTCAGGGGG<br>GCCTGAGGAGG<br>CCGACTCCTC   |                              | ACCETE AGAC<br>TOGO AGTETO<br>CCCCGTGACE<br>GGGGCACTGG  | OPCTACTCCC<br>CAGATGAGGG  | CCACCCITCAG<br>GGTGGCAGTC                    |  | ACCICCIOCC<br>TOCCCCCCO<br>CACOCCCCCA<br>GTCCTCCOOT    |                              | -  | CACACCCGITG<br>GICCCCGITG<br>GACTCGTACT<br>CTGAGCATGA        |
| TCTQQQQQT AGACCTCCCA ANCAGCAGTG TTGTCGTCAC AGCGGCCCCA AGCGGTGTCAC AGCGGTGTGCCACG AGCGGTGTGCCACG AGCGGTGTGCCACG AGCGGTGTGCCACG AGCGGTGTGCCACG AGCGGTGTGCCACG AGCGGTGTGCCACG AGCGGTGTGCCACG AGCGGTGTGCCACG AGCGGTGTGCCACG AGCGGTGTGCCACGCACG AGCGGTGTGCCACGCACGCACGCACGCACGCACGCACGCA  |                              | AACTACOOCO<br>TTUATOCCÓC<br>TOATOCAAGA<br>ACTACOTTCT  | COACCAGGCC                | CCCACCATCA                                   |  | CCCGGGCACT<br>GCGCCCGTGA<br>ACTACACGCC<br>TGATGTGCGG   |                              | 3  | CTCAGGGTCG GAGTCCCAGC ANACTACTA TTTGATGAAT                   |
| CCTACGATGA<br>CGANTICTACT<br>ACCCCCCTCC<br>TCCCCCTCC<br>CCCCTCTCC<br>CCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTC  | AGNAGAAACC<br>TCTTCTTTGG     | CCFTGCATAC<br>GGAACGTATG<br>TTGCCAGACA<br>AACGGTCTGT  | COANGATOTT                | CCCGCCAGCC                                   | CACTOGFAAT<br>GCATGFCCAT<br>COTACAGGTA | AOTOCOCOTO<br>TEACOCOCOC<br>GAGGEGEGEA<br>CTECOCOCO    | GACGGCTCAG                   | ACOOOCOCC<br>TGCCCOCCOO                  | AGTOCTATOA<br>TCACGATACT<br>TTGCAAGAAA<br>AACGTTCTTT         |
| 14501  | 14801                        | 14901   | 15101                     | 15201  | 15401                                  | 15501  | 15701                        | 15801                                    | 15901  |

Figure 15 J

# PMRKAdSqag MER682

| 16101 | CCAGGTCATC        |               | TCTATEOCC        | CCCGAAGAAG                      |                   | ATTACAAGC                  | CTCAAACCTA        |                          | ANAMOMANA          | GNANGATRAT                              |
|-------|-------------------|---------------|------------------|---------------------------------|-------------------|----------------------------|-------------------|--------------------------|--------------------|---|
|       | COTCCACTAG        |               | AGATACCOOG       | COCCCCTCT AGATACCCOC GOGCTTCTTC | בשות שונים שנים   | דאאזידדיניגן קאיכידידינכאד | GCCTTTCGAT        | CCCAGT                   | माराजाम            | כנדדנכדאכדי                             |
| -     |                   |               |                  |                                 |                   |                            |                   | Safi                     |                    |   |
| 16201 | CATCATCAAC        | TTGACGACGA    | CONTRAMETO       | CTRCACCCTA                      | ניגעטערינינאני    | <b>ATDINUTATION</b>        | CACTOGAMG         | GITCGACCCCT              | AAAACONOTT         | PTGCGACCC"                              |
|       | CTACTACTTO        | AACTGCTGCT    | CCACCTTGAC       | GACCITOCGAT                     | נאשנטנונטנונונו:  | COCTOCCC AT                | GICACCTITIC       | CAGCTGCCCA               | TTTTOCACAA         | AACCCTGG:::                             |
| 16301 | GCACCACCOT        | AGRETITACO    | CCCCCTTAGC       | CONCOCCO                        | CACCITACAAG       | CACHINITAIN                | ATGAGGTVTA        | CHOCCACTIAN              | CACCTOCTTO         | ACCAGGCCAA                              |
|       | COTOCTOCCA        | TCAGANATGC    | <b>CACCACTOR</b> | CGARGTTAGA                      | CHUCATOTTC        | CCTICACATAC                | TACTCCACAT        | OCCOCHOCITO              | CTCGACGAAC         | TCCTCCGGTT                              |
| 16401 | COMIXECUCAL       | COCCACTITO    | CCTACGGAAA       | REGERATANG                      | CACATISCING       | CONTRACTION                | CCACCACKCC        | ANCCCANCAC               | CTAGCCTAAA         | OCCCGTAACA                              |
|       | OCTOCOCOAO        | CCCCTCAAC     | COATCCCTTT       | COCCOTATIC                      | CTGTACGACC        | GCAACOGCGA                 | ccrecrecce        | Trecerrore               | GATCOGATIT         | COCOCATTOT                              |
|       | P.A.              |               |                  |                                 |                   |                            |                   |                          |                    | Kim                                     |
| 10591 | GACOTCOTCC        | ACGACGCCCG    | CCAACGTCGC       | ACCCAACAAA<br>ACCCTTCTTT        | ACCICCCCC T       | AAAGT CECCAG<br>TTTCCCCCTC | AGACCACTOA        | TOCCACCCAC<br>ACCGTOCOTO | CONCEASONS GCACONC | AMMOTACC.<br>TACCATORIGT                |
| 16601 | ACCOCCAGCG        | ACTOGRAGAT    | GTCTTGGAAA       | MATGACCGT                       | COAACCTRA         | CHREMACCCG                 | AGGTCCGCGT        | GCGCCCAATC               | ANGCAGGTGG         | COCCGGGACT                              |
|       | <b>PCOCOGICOC</b> | TOACCTICTA    | CAGAACCITT       | TTTACTOCCA                      | CCTTVIGACIC       | <b>GACCTCGGGC</b>          | TCCARGCGCA        | CCCCCOTTAG               | TTCOTCCACC         | <b>GCGGCCCTGA</b>                       |
| 16701 | OCCUTOCAG         | ACCOTOGACG    | TTCAGATACC       | CACTACCART                      | AGCACCANTA        | THECCACCGC                 | CACAGAGGGC        | ATOGAGACAC               | AMCGICCCC          | GOTTGCCTCA                              |
|       | CCCGCACOTC        | TOCCACCTOC    | ANOTICTATOR      | GTCATORITCA                     | TCUTGGTCAT        | AACIGOTOGCO                | ororcicco         | TACCTCTOTO               | THYOCAGOOG         | CCAACCOACT                              |
| 16801 | accontacco        | ATOCCOCOOF    | OCACCCCCT        | <b>GCTISTOGGCG</b>              | CONCENSAC         | CTCTACHAG                  | GTTSCAAACTO       | ACCCUTODAT               | GTTICGCGTT         | TCAGCCCCC                               |
|       | COCCACCOCC        | TACOOCUCCA    | CONCCIOCONO      | כמשכמכבמנוב                     | <b>GCARTITUTE</b> | GAGATACCITC                | CACGITTICACC      | TOGGCACCTA               | CANAGCOCAA         | ACTCCCCCCC                              |
| 16901 | 000000000         | CCOFTCGAGG    | Addracticco      | CCCCCAGCGC                      | <b>GCTACTGCCC</b> | ONTATRICCC                 | TACATCCTTC        | CATTCCCCCT               | ACCCCCGGCT         | ATCGTCGCT                               |
|       | CCOCOOCOC         | OCCARGETEC    | THEATGCCOS       | <b>OCCOSTICICO</b>              | CCATTANCOO        | CTTATACGGG                 | ATCTACCAAG        | GTAACGCGGA               | TOGGGGCCGA         | TAGCACCGAT                              |
| 17001 | CACCTACCOC        | CCCAGAAGAC    | GAGCAACTAC       | CCGACGCCGA                      | ACCACCACTG        | CANCECCCC                  | CCCASCIFICOGC     | CCTCGCCAGC               | CCGTOCTAGC         | CCCGATTTCC                              |
|       | <b>OTGGATGGCG</b> | <b>COUNTY</b> | CTCGTTCATC       | <b>GOCTGCGGCT</b>               | TROTOGRAC         | CTTONGCOGC                 | <b>GOCKERVOCO</b> | OCAGCOGICG               | COCACGACCO         | GGCTANAGC                               |
| 17101 | CTCCCCAGGG        | TOOCTCOCGA    | ACHIAGCICAGG     | ACCCITAGING                     | PRECNENIC         | PROCTACCAC                 | CCCAGGATATCG      |                          | concriniona        | GTTCTTGCAG                              |
|       | CACGCOTCCC        | ACCOAGCCCT    | Techeconee       | TRACACCACO                      | ACCEPTIGACE       | CACGATGGTO                 | <b>GGCTCGTAGC</b> | AAATTTTCOO               | CCAGAAACAC         | CAAGAALGTY                              |
|       |                   |               |                  |                                 |                   |                            |                   |                          |                    | T. S.                                   |
| 17201 | ATATOOCCCT        | CACCTOCCOC    | creedmine        | COCTCCCOCG                      | ATTECCANOGA       | AGANTGCACC                 | GTAGGAGGGG        |                          | CACOOCCIOA         | COCCURRANT                              |
|       | TATACCOCCA        | GTOGACGGCG    | GAGGCAVAGG       | <b>GCCACAGCCC</b>               | TAMASICTOCT       | TRITACGINGO                | CATCCTCCC         | GTACCOGCCG               | OTIGE CE COORCE    | <b>OCCCCCCOTA</b>                       |
|       | Ē                 |               |                  |                                 | Sphi              |                            |                   |                          |                    |   |
| 17301 | aconcorac         | CACCACCOGC    | <b>GOCCOCCCC</b> | GTCGCACCGT                      | CYN, ATVICACG     | <b>GCGGTATCCT</b>          | DCCCCTCCTT        | ATTCCACTOR               | receeses           | האדדשמכמה                               |
| ]<br> | COCADCACGC        | orcorocco     | CCCCCCCCCCC      | CAGCGTGGCA                      | <b>GCCTACGCVC</b> | CONCATARA                  | CCCCCCCCCC        |                          | AGCGGCGCCC         | CTAACCCCC                               |
| 17401 | GTOCCCOGAA        | TTOCATCCOT    | GGCCTTGCAO       | GCOCAGAGAC                      | ACTICATTANA       | AACAAGTTIAC                | ATCTINGUARA       | -                        | AAAGTCTGGA         | CICICACGIT                              |
| ı     | CACOOOCCTT        | AACCTAGGCA    | CCCGANCETC       | COCONCION                       | TOACTAATT         | THEFTEMED                  | TACACCTITIT       | TAGITITATI               | TTTCAGACCT         | GAGAGTCCCA                              |
|       |                   |               |                  |                                 |                   |                            |                   |                          |                    | A 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 |
| 17501 | COCTIOORCC        | TOTANCTATT    |                  | GANGACATCA                      |                   | הנידטיבניכט המאכאהמסכי     |                   | COCOCCOOL                | CATCOCIAAAC        | TRACEMENTA                              |
|       | <b>GCGAACCAGG</b> | ACATTGATAA    | MCATCTTAC        | CENCICANGE                      | TGAAACGCAG        | AGACCTARCA                 | OCTOTOCCCA (      | GCGCCGGGCAA              | GFALCUTING         | ALCO 15 180                             |

Figure 15K

# PMRKANISqag MRR682

|       |               |                   |                   |                          |                   |                  |                  |              |                    | CENTER MICE.            |
|-------|---------------|-------------------|-------------------|--------------------------|-------------------|------------------|------------------|--------------|--------------------|-------------------------|
| 7601  | PCOCACCAG     |                   | Gritoreceret      |                          |                   | ACCESSIVATIVA    | DELUTE AND COLOR | ANGETOCCAA   | TETTGATAC          | CGTCGTTCC               |
|       | AGCCOTCOTC    | CTTATACTCO        | CCACCGCCGA        | METH GACTUC              |                   |                  |                  |              |                    | i alicatorio de de      |
| 10771 | CTOGAACAGC    | MOCACAGAICC       | AGATTSCTTANS      | CCATAACTTO               | _                 |                  |                  |              | CUICIGOCAI         | Marchaelle.             |
|       | gacctronco    | recharces         | TCTACGACTC        | CCTATTCAAC               | THE TANKE THE     | TANAGGTTGT       | THECACCAT        | CTACCOGACC   | GGMG/ACCOIN        |                         |
|       |               |                   |                   |                          | - Herefill        |                  |                  |              |                    |                         |
| 17803 | GTOGACCTOD    | CCAACCAGGC        | ACTGCANAT         | ANGATTANGA               | GTANZITICA        |                  | CCCATTACACIO     | AGCCACCACC   | COCCUTOURO         | ACAGIGICT.              |
|       | CACCTGGACC    |                   | TCACGITTITA       | TTCTAATTGT               | CATTETIAACT       | ACKYNGCOMIN      | CHANCA TO THOS   | rccanactoo   | CCGGCACCTC         | TOTCACAGY               |
| 1001  | CAGAGGGGCG    |                   | CONTICACACIC      | CCGACARAGA               | AGAMCTOTA         | TTV:ACTOCAAA     | TAIACGAGGC       | recereative  | CHOCAGGCAC         | TAMATETANT:             |
|       | GIVENCECO     |                   | GCAGGCGCG         | GRETETECET               | TCTTTTACAC        | CACTOROTTE       | ATCTGCTCGG       | AGGGAGCATG   | CTCCTCCGTO         | ATTICETIC:              |
| 1000  |               |                   | transment.        | CCCTACTOGA               | משלאן לאנט        | NATACACC         | COTAACCCTC       | GACCTGCCTC   | CCCCCCCCCA         | CACCCCACCA:             |
| 10001 | COACGOOTOO    |                   | AGCCCCCCTA        | AGCOCOCOTA . CCGATYRICCT | כאכניאכניכניפ     |                  | GCATTGCGAC       | CTOGACOGAG   | GGGGGCGGCT<br>Pref | GIVIDOTCOD:             |
|       |               |                   |                   | CHARLES M.C.C.           | لتهديهاوردو       | כמכנונטבום       | כפככפכפיכפ       | CCAGCIXOTCC  | OCCUNICATIO        | COCCCOTAG               |
| 10101 | MACCIFICAC    | ACCURACECT        | CHECCECOTA        | CAACATIVAG               | CACCIATICIATIC    |                  | מכסטממטמטט       | CONCOCCAGO   | COCTAGCAAC         | GCCCCCCCATC             |
| .0.0  |               |                   | ACACTORACA        | GCATCGTGG                |                   | CANTCECTEM       | AGCCCCGACG       | ATTOCTTCTOA  | TAGCTAACOT         | GICOTATOR               |
| 10201 | GENCACCUTE    |                   | TOTOACTION        | COTAGCACCC               | AGACCCCCAC        | GITTACCISACT     | TEGEOGETINE      | TACCANGACT   | ATCGATTOCA         | CACCATACA               |
| 10101 | TATICATUTAT   | _                 | COCCOCCAGA        | ממשכתפכת                 | ACCICCURE         | CIRCCIOCTIT      | CCAAGATOSC       | TACCCCTTCO   | ATGATGCCCC         | ACTICOTOTION            |
|       | ACAGTACATA    | _                 | GCGGCGGTCT        | CCTCGACGAC               | TCCACCACCC        | מכננסטנמאא       | CAGITICTACCO     | ATGGGGAAGE   | TACTACGGCO         | ירא באטע                |
| 18401 | CATGCACATC    | 1CGGGCCAGG        | ACCCTCGGA         | GTACCTGACC               | CCCGGGCTGG        | TECNOTTING       | כבמבמככענב       | GACACOTACT   | TCAGCCTGAA         | TANCARGITT              |
|       | GTACOTOTAG    | -                 | TOCOGAGCCT        | CATCHIACTEG              | <b>CHECCCGACC</b> | ACCITCAAACO      | CCCCCCCCCCTCC    | CTCTGCATGA   | AUTOMACIT          |                         |
| 18501 | MGAMACCCCA    | _                 | TACCCACGAC        | GTGACCACAG               | ACCIMENTER        | CCCTTTCACO       | CTCCCATTTCA      | recentration | CCOTCAGGAT         | MCMCCGTMC.              |
|       | TCTFTGGGGT    | -                 | ATGCOTGCTG        | CACTURENTIC              | TOGCCAGGGT        | המכתאהוה         | CACCCCAAGT       | AGGGACACCT   | GGC ACTC IN        | Nancy Control           |
| 18601 | COTACAAGGC    | <b>BCOGFFCACC</b> | CTAGCTV:TGG       | CTGATAACCT               | TOTACTODAC        | ATGUCTTCCA       | COTACTITICA      | CATCCOCOGC   | GIGCTGBACA         | GOOT CLIN               |
|       | OCATOTICCO    |                   | <b>GATCOACACC</b> | CACTATTCCC               | ACACGACCTG        | TACCGMOST        | GCATGAMCT        | GIAGULGELO   | בערכאררופו         |                         |
| 18701 | TTTAAGCCC     |                   | CTCCCTACAA        |                          | CCCAAGGGTG        | CCCCAAATCC       | TYGCGAATGO       | GATCAAGCTG   | CTACTGCTCT         | TGAONTAON<br>ACTUTANTUM |
| !     | AAAATTCGGG    | ATCIACACCOT       | GACOGATGTT        | <b>GCGGGACCTIA</b>       | ממטבבאנאנאני      | CAXICTITAGE      | MCGCTINCC        |              | Carlo Partie       | COTATABATA              |
| 18801 | CTAGAAGAAG    | ADDACCIATOR (     | CAACCAAGAC        | _                        | ACT TANKE TOA     | GUNGCAAAAA       | ACTCACGTAT       |              | _                  | CCATATTTAT              |
|       | GATCTICITIC   | TCCTGCTACT        | gricencia         | _                        | regrector         | COLCULATION      | Niversity And    |              |                    | CTCAGTV3GTA             |
| 18901 | TTACAAAGGA    |                   |                   | -                        | ACCTANATAT        | GCCGATAAAA       | CATTICACE        |              | -                  | GACTICACCAT             |
|       | AATOTITICCT   | · CCCATAAGTT      | TATCCACAGC        | TICCAGITING              | TGGATTTATA        | CONCINE          | SINGE IN         | -            | A A CONTRACTOR A   | TURARATES!              |
| 19001 | COMMICAGIA    |                   |                   |                          | ANGACTACCC        | CAATGAAACC       | TACAATTE         | AGTATACGTT   | THOODIETT          | ACTITITACE"             |
|       | OCTITION      | TAATTAGTAC        | GICGACCCIC        |                          |                   |                  |                  |              | ACCORCAGGC         | AATGGTGAT .             |
| 19101 | COCCARCCCA    | -                 |                   |                          |                   | CHANNATCH AM     | MANAGAGIT        | GATICACTICCO | TCGCCGTCCO         | TTACCACTAI              |
|       | cccorrector   | PAGAACATTT        | _                 | -                        |                   | יביים בו שנייונו | ATATEM TTA       |              | ATTAMBANG          | <b>OTAACTCAC!</b>       |
| 19201 | ACTITICACTICC |                   | -                 |                          | TATA MANALC.      | CATHUTCH GAG     | TATAMGANT        |              |                    | CATTGNOTO:              |
|       | TCAACTCAGG    | 3 ATTHCACCAT      | AACATGTCAC        | TACTACATO                |                   |                  |                  |              |                    |                         |

Figure 15L

| CCCATTATAC<br>CCCATTATAC<br>MOSTANCIAN<br>ANDITACONA<br>TTGAN SITT   | AACATAGES: TTGTATCGS: TGGCTCCCS: AGCGAGGGS: ACCGAGGGGS:  | CCTGCGCTAT<br>GGACGCGATA<br>CCGGGCTCAT<br>GGCCCGAGTA   | ROTTTGATAL:<br>ICAAACTATC:<br>CTTTAACO! :<br>GAAATTGC! : | CCCCGACTORY<br>GCGCCGACTY<br>ATGGAACCTTY<br>TACCTTYGAAATT | CAMCTITAA<br>ATTRICTACS<br>TAACCGATGG                   | GATTTATOTT<br>CCCTRCTAN:<br>GCCACGATTO               | TCCARTAACT<br>AGGTCATTOA |
|--|--|--|--|---|---|--|--------------------------|
|  | CCCTTTTCT CC CCCTGGACATONOG CGACATONOG CACATONOG CACATONOG CACATONOG CACACONOG OCAMPECTOS COTTACGACC CCTTCTCCTO GOAAGAGGAC GOAAGAGAGAC GOAAGAGAGAC GOAAGAGAGAC GOAAGAGAGAGAC GOAAGAGAGAC GOAAGAGAGAC GOAAGAGAGAC GOAAGAGACCAG GOAAGAGACCAG GOAAGAGAGAC GOAAGAGACCAG GOAAGAGAGAC GOAAGAGACCAG GOAAGAGACCAG GOAAGAGAGAC GOAAGAGACAG GOAAGAGACAG GOAAGAGACCAG GOAAGAGACAG GOAAGAGACAG GOAAGAGACAG GOAAGAGAGACAG GOAAGAGAGACAG GOAAGAGAGACAG GOAAGAGAGAGACAG GOAAGAGAGAGAC GOAAGAGAGAGAC GOAAGAGAGAGAGAGACAG GOAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG |  | CCCCCAACC CCCCCAACC CCCCCAACC CCCCCAACC CCCCCAACCA        |   |  | CATCCCATTC               |
| CTAATCHAT CATACCAGCT GTATGOTCGA AMACATGGA TATAGTACCT GAAAATGGAT  | CTTTTACCTA<br>GGAGAAATTT<br>CCTCTTTAAA<br>CTACATGAAC   | AACCACCACC<br>TTOOTOOTOO<br>TTAAAAACCT<br>AATTTTTGGA   |  | • • • •   |   | GCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG               | CCCTTFGGCG               |
| TTTATTCH<br>ANANTANGGA<br>AGAGAGETT<br>TSTCTCGAAA<br>GANTATTCA<br>CTTANTAAGT   | THORICONITY<br>GCCANCCTIFT<br>CGGTTGGACA<br>ACACCTACGA   | CAACCCATT<br>GTTGGGTAAA<br>TTCTTTGCA   | <b>~.</b> • •  | TCCATCCCCT AGGTAGGGGA ACACCTACTC TGTGGATGAG               | TESCHATCAC ACCGTTACTG TTCCTFEFTAC AAGGACCATG            | AGCCCATANG<br>TCGGGTACTC<br>TCCCCCACC<br>ACGCGGGTTGG | TYCGCTAGCGT              |
| TTATACH MA<br>ANTECETITE<br>ACACAMANE<br>TENTICITITE<br>ECAGANETTA<br>GATETACAMT   | ATTTEGATT CANTCTAAAT GTTAGATTTA GATAACCCAA   | TYTACAACGT<br>ACCTGTTGGA<br>GCCTCAGAAG<br>CGGAGTCTTC   | TCCCTAGGAA<br>AGGATCCTT<br>CGCTTGAGGC<br>GCGAACTCCG      | CCTGCCCATA<br>GCACGGGTAT<br>GACCCTTATT<br>CTGGGAATAA      | TCARCTCACC<br>ANTCAACCO<br>CAAAAACTTC<br>GTTTCTGACC     | AGAAACITCC<br>TCTTTGAAGG<br>TTGGCTACCT<br>AACCGATGGA | ANOTTHEFT<br>TITEMMENA   |
| TACATTICAN<br>ATGATTICAN<br>TAGATTICAN<br>ATCTANACTI<br>CAGGTANTIAT<br>GPCGATACTAN   | CCATCHARTEC<br>CCATCHTAAT<br>GGTACCTTA<br>AAAAATTECT<br>TEEFTAAAAGA  | CTTGACTATA<br>GAACTGATAT<br>ACATTCAAGT<br>TGTAAGTCCA   | TCTTFTAGAGC<br>ACACATCTCG<br>ACACACCTCCA<br>TCXXCXICGGT  | NCCCTACCAA<br>TCCGATGGTT<br>CTCCGGCTAC<br>GAGCCCGATT      | CACTCTTCTY CTCAAAAAAC CTTAAACATCAAC CATTGTACTO          | CTCCTTCTTT<br>GAGGAAGAAA<br>TCTGGATTTG<br>AGACCTAAAC | TTACCCAGNA<br>AATGGGTUTT |
| CAGGCCTANT GTCCVFATTA AATTACTICTTU FFACGACAAC TCCCACACACCT TACAGAGACT TACAGAGACT   | ANTANTITIC<br>TTATTANAC<br>CTTCCAACGT  | ACGCTCGTCC TGCGACCAGG GTGCCCTTCC CACGGGAAGG  | TTANCATOST<br>AATTOTACCA<br>GOCCCACAAC<br>CCGOSTUTTS     | ATACCCGCCA<br>TATGGGCGGT<br>CATCACTGGG<br>GTAGTGACCC      | CATTACCTTT<br>GTAATGGAAA<br>GTTGCCCAGT<br>CAACGGGTCA    | ACCCCATGTA<br>TROCOTACAT<br>ACACAACAAC               | GATGACAGCA               |
| CPATRICICAA<br>OATACOSCITT<br>ARCCISAITTS<br>TACCISAARC<br>TACACCITAA  | AAGAGTTKGA<br>TTCTCAACCT<br>AAGTACAGTC   | ACCTTGGAAC<br>TGGACCTCG<br>TGGTCGCTAT<br>ACCAGCGATA  | AGGANGZATO TECTTECTAC "TETTECECAT AGAAGGGTA              | GCTCTACCCT<br>CGAGATGGGA<br>AAGGAAAGCC<br>FTCCTFTGGG      | AGANGATIGOC<br>TETTECEACEG<br>GGGTTACAAC<br>CCCAATRATIG | AGCTACAAGG<br>TCGATOTTCC<br>TCCTACACCA<br>AGGATGTGGT | CAAGACCTATA              |
| GGCCAACAAT CCGGTGTTA CGGCCGGTTCTA GCCCGGTTCTA GCCCGGTTCC GCCCGGTTCC GCCCGGTTCC CCCCGGTTCC CCCCCGGTTCC CCCCGGTTCC CCCCCGGTTCC CCCCCGGTTCC CCCCCGGTTCC CCCCCGGTTCC CCCCCCGGTTCC CCCCCGGTTCC CCCCCCGGTTCC CCCCCCGGTTCC CCCCCCGGTTCC CCCCCCGGTTCC CCCCCCCGGTTCC CCCCCCGGTTCC CCCCCCCC  | AMANTGAMT<br>TTTTACTTTA<br>CCACAAGCTA  | TOCTACATTA<br>ACGATGTAAT<br>TOCTGGGCAA<br>ACGACCCGTT   | OTOGAACTIC<br>CACCTTGAAG<br>TACCCCACCT<br>ATCCOOTGGA     | CCCCCAACAT<br>GOCGOTTOTA<br>CCTTAAGACT<br>GGAATTCTGA      | CACACCTITA<br>OTOTGGGAAAT<br>TIGACGGGGA<br>AACTGCCCCT   | TATCCAGAG<br>ATAGGGGCTC<br>CAGGTGGGCA<br>GTCCACCCGT  | CCCTTATAGG               |
| AGAACTANTO COTOTTCOO COTOTTCOO COTOTTCOO COTOTTCOO TATAGAACCOO GAACTATAGAACTAGAACTATAGAACTATAGAACTAGAACTATAGAACTAGAACTATAGAACTAGAACTATAGAACT | AATGACGAA<br>TTTTCAGATA<br>AAAAGTCTAT<br>TGTATTTGCC  | OCTADTOGAC<br>COATCACCTO<br>COCTCAATOT<br>GCOAGTTACA   | ACACCTACGA<br>TOTOGATOCT<br>CATTTGCCTT                   | TATCTCTCCC<br>ATAGAGAGCC<br>CCTTCACGCG<br>GGAAGTGCGC      | TTACCTCANC<br>AATGGAGTTG<br>AAGGGCTCAG<br>TTGGCGAGTC    | AGGCTTCTA<br>TCCCGAAGAT<br>GGACTACCAA<br>CCTGATGGTT  | TTCCCCTATC<br>AACCCCTATC |
| 19301<br>19401<br>19501  | 19701  | 19901  | 20201  | 20301   | 20501   | 20701  | 20901                    |

Figure ISM

| •      | CCATOGALGA                                    | CACGCCCTT ' GTGCGGGA 1       | TCAAAGATCT<br>AGTTTCTAGA  | ODCCOOTCY > 1                   | COACTCAAGC               | AAAGCGTAC .<br>TTTCGCATGT | CACCATOAN' GTGGTACTT     | OAGCGCCAC"<br>CTCGCOOTO | CTTTCAATAA               | CCCATCCCTA               | TCACTCCAC!                      | ONTACACACA<br>CTANTICITY | GTTGCTCAGO               | TGACCOTCC             | CGTANGACTT                |                     | OGICCCACCO OFFICIPICACO CCOCOCTUCO CAACAAGTGC           |
|--------|---|------------------------------|---------------------------|---------------------------------|--------------------------|---------------------------|--------------------------|-------------------------|--------------------------|--------------------------|---------------------------------|--------------------------|--------------------------|-----------------------|---------------------------|---------------------|---|
| Randil | GNOSTICIATE CENTOGRICA CTCCACCTACT GGIACCTGCT | TOTACCTOCO<br>AC/ TOCACOC    | AAAGCCATTO                | TAGTCAATAC                      | TTCTCACCAG<br>AAGACTCGTC | AAGTCCACCC<br>TTCAGGTGGG  | ATCACAACCC<br>TAGTGTTQQQ | CAGCTÍCCTO              | ACTAGAMACA<br>TOATCICTOT |                          | CCACTICAAA                      | CGCGAGTFGC               |                          | CATCAAAAGG            |                           |                     | GGCCCCACCG  |
|        | CATCACTETT                                    | ATCGAAACCG<br>TAGCTTTGGC     | OCAGGAACTG<br>COTCCTTGAC  | <b>GCCTGCGCCA</b><br>CGGACGCGGT | CCTTROGCTT               | AACOCTOGAA                | ACTCCCATED<br>TGAGGGTACC | AACAGCTCTA              | AAAATAATOT               |                          | COCCOTCCAG                      | gcccracaca               |                          |                       |                           | -                   | ACCACATTIC<br>TOSTCTANAG                                |
|        | ACCCCCTAGA<br>TOCCCCAATUT                     | היהנונמאריות<br>מווכאככטו:Ad | GCTCCANTOA<br>COAROTICACT | ACACAMOCTC<br>TSTGTTCGAG        | CTCTPTOAGE<br>GAGAMCTEG  | ACCOCTGTAT<br>TGGCGACATA  | CTCCCCCCAA               | COCMCCAGO<br>CCGTTGGTXC | ANAACATOTA<br>TTTTGTACAT |                          | ACAACCATCC<br>TOTTGGTAGG        |                          |                          |                       |                           | CTCCGAAACG          | CCTCTAGACO  |
|        | אארידגיניניני<br>דיוניאמיציניזים              | ACCAMPOCACA<br>TUGITORICOCOT | GUCCICATOG                | THEFFICTEC                      | AACATGCTAC<br>TTGTACGATG | TCTTCCCCCG                | CCTTTRCCAA<br>GGAAACGGTT | CACCCTACGE              | TOTCACTTGA<br>ACAGTGAACT | THACCGRETT               | AAACTCAGGC<br>TTTGAGTCCG        |                          |                          |                       | AMAGCAACT                 | TTTCGGTGGA          | OFCOTOTIONS CARCACTTO CONCROTOTIC CAGGACTACA            |
| •      | TCTCTACOR C                                   | GAGGGAGAGA                   | AACAACAGET<br>TIGITIGICIA | TTTCCAGGTT<br>AAAGGTCCC3A       | COCACTICAAA              | CCCCATTCCT                | AMBAGGGGG<br>AMBAGGGGGC  | AGGTACACC<br>TCCATGTCGG | CACTICITIT               | ACCCCCACCC               | ACCACACTA<br>ACCACACATGAA       |                          | CETATAGAAC               |                       | COCCACACORG<br>GATCTGCTTA | CTAGACGAAT          | GREGINEAND CAREACETTS CAGEACOTTS CAGEACOTTS GTECTTSGAAC |
|        | CCCAAAACCT                                    | CHTTGACHG                    | AACCAACATC                | TENCANGERE                      | CCCACCTION               | TCCCCCCTAG<br>ACCCGCCATC  | CHECTECATO               | AACAGTCCCC              | TTAGGMGCGC               | GTGATTATIT<br>CACTANTANA | TOOTOTTTAG<br>ACCACAAATC        | GOTCOOCCC                | CCAGCCCGCG               | CACCACGTGC            | TAMARICETT                | ATTITICGGAA         |   |
|        | ACAGACCTGG<br>TGACTGGACC                      |                              | ATMAGGAAGC                | TECCCALCTA<br>ACCCCTCGAT        | CTACCOMAA                | CTCAGTGAGG                | OTOCACTATT<br>CACCTGATAA | CTCCATGCTC              | AGTOCOCACA               | ACACTCTCGG               | OTTOCCATAC<br>CAACGCTATG        | GCGTTTAGCA               | COCMATCOT                | ACTCCCCCCC TAGCTCCCTT | ATCCACOCAA<br>AGCCCCTTACA | TCCCCGGACGT<br>Stil | ONCACASCEGE<br>CTGTCCGGCG                               |
|        | GOGCICACTC                                    |                              | ACCCCACAAC                | CCATATETT                       | COCATOTORC               | CANACTCATO                | TCOCCOCCT                | Kont<br>GOOTACCCAA      | CCCCAOCCAC               | TTTTATTTOT<br>AMATANCA   | <b>GCAGGGACAC</b><br>COTCCCTOTG | CATCACCAAC               | OTACTOOTTO<br>TOGAACACTA | ACCTIGIDAT            | AGTTAGGATAC               | CAATCCTATG          | TGATTCGCCG<br>ACTAACCGGC                                |
|        | TTATOTCCAT                                    |                              | TYCOCTOOCA                | 1001101000                      | CACACACCCC               | AGGITTACCA                | CCCCOOUTS                |                         | COCCCTACT                | ACCOUNTACT               | TOCOCCACTO<br>ACOCCOTOAC        | OCCTOCOCAC               | CCGACGCGTG               | CAACOTCGTO            | CUCTIOCCTC                | GCCAGACCC           | <b>GCCOGNAAAC</b><br>CGCCTTTTO                          |
|        | 21001   | 21101                        | 21201                     | 21301                           | 21401                    | 21501                     | 21601                    | 21701                   | 21801                    | 21901                    | 22001                           | 22101                    | 10226                    | 22101                 | 22401                     |                     | 22501   |

Figure 15N

# PNRKAUSgay MER682

| 22601 | ATCTICOCCT   | ATCTTOSCCT TRICTAGACTE | CTRICTATE   | כידיה יודיאמב ממתמאבוריה                                 | COMPRESS CONCACATO |              | ATTICAATICA       | CONOCICCIT        |                                 | ATGCTTCCGT          |
|-------|--|------------------------|-------------|--|--------------------|--------------|-------------------|-------------------|---------------------------------|---------------------|
| <br>  | TAGAACCIACA  | TAGARCCICA ACGARCTICAC | GAGISANTICO | פאניונאאראובים בעכנשיפאנים עכאאאטנינא                    | CCANAGGGA          | CCACTOTACE . | TANACITAGE        | GCACGAGGAA        | TAAATAGTAT                      | TACGNAGGICA         |
|       |  |                        |             |  |                    |              |                   |                   |                                 | <u>.</u>            |
| 22701 | GTAGACACTT   | AAGCTCGCT              |             | TCCATT TCAG CGCAGCGGTG                                   | CNACTACAAC         | שלמלאשככנום  | recerete          | ATCCTTCTAG        | GICACCINCING                    | CNANCOACTO          |
|       | CATCTOTOM  | TTCGAGCGGA             | AGCTACACTIC |  | CHANNETE           |              | ACCEGARACAC       | TACGAACATC        | CAGTOGAGAC                      | CTTTY CTKIN.        |
|       | Pstt   |                        |             |  |                    |              |                   |                   |                                 |                     |
| 22801 | CAGGEACGCC   | TOCAGGAATC             | OCCUCATICAT | CCTCACAAAG   | GETTATAGE          | TODANTON     | CAGCITACIAC       | ccocconact        | CCTCOTTICAG                     | CCAGGICT" . I       |
|       | GTCCATGCCC   |                        |             | <b>CCANTOTTTC</b>  | CAGAACAACG         | ACCACTTORA   | GICCOACGING       | GOCCCCACCA        | GGAGCAAGTC                      | <b>GTTCCACIANC</b>  |
| 22901 | CATACTACK  | CCAGAGCTIC             |             | CACTIGGTCA GOCAGINGTT                                    | TOAAGTTOSC         | CTTTAGATEG   | TRATCCACGE        | GOTACTIGIC        | CATCAGCGCG                      | COCOCAGOLIF         |
|       | GTATOCCOCC   |                        |             | CCOTCATCAN ACTICAAGCG                                    |                    | GAAATCTAGC   | <b>NATAGGTGCA</b> | CCATGAACAG        | GTAGTCGCGC                      | acacatcar 1         |
|       |  |                        | Fred        |  |                    |              |                   |                   |                                 |                     |
| 23001 | CCATGCCCTT   | CTCCCACGCA             |             | GACACGATUG GCACACTCAG CGRGTTCATY:                        | CGGGGTTCATK:       | ACCOUNTER    | CACTITICCEC       | THEOGRAPHICA      | TCTTCCTCT.                      | CCICINGO            |
|       | OUTACOODAA   |                        | CTGTGCTAGC  | COTOTOACTO   | GCCCAAGTAG         | TYCCATTANA   | GTGAAAGGGG        | AAGCGACCCG        | AGAAGGAGAA                      | COCAGONACGCA        |
| 21101 | CONTATACCA   | CTACACCACTO            | CONCONCINC  | ATTICACKCICAC  | FUCACTOTOC         | GCTTACCTCC   | TTTCCATCC         | TTCATTAGCA        | ccoorcoon                       | <b>OCTUMAN</b> CING |
|       | OCCURATOR  | _                      |             | TANGTERACT   | GCGTGACACG         | CONTRACTOR   | ANACGCTACG        | AACTAATCGT        | <b>GGCCACCCAA</b>               | CONCILITION         |
| 27201 | ACCATTRATA   | _                      | PRETETACE   | <b>PCCPNIXITAT</b>                                       | CCACGATTAC         | CICHICITATI  | GREGGEGET         | COCICTICOC        | ACANGGGCGC                      | THETHERE            |
|       | TCCTAAACAT   | _                      |             | _  | ORTGETANTO         | GNONCCACTA   | CCCCCCCCCA        | GCCCOANCCC        | 1ctrcccoco                      | ANGAANNGA           |
| 21301 | Tr. T. Transferde  | -                      | 1ccaccacc   | ACCITICATING   | CCACACACTO         | CONTRACCCO   | GCACCAGCGC        | <b>GICTIOTOAT</b> | GAGICITICET                     | CONCCINCOCA         |
|       | AGAACCCGCG   | _                      | AGGCGGCGCC  | TCCAGCTACC   | GOCGCCCCGAC        | CCACACCCCC   | COTOCINCOCO       | CACIAACACTA       | CTCAGAAGGA                      | <b>GCAGATAGCCT</b>  |
| 21401 | CHATACOC   | _                      | -           | 000000000  | BRARDCRACG         | CCCACCOGGA   | CHARGACGAC        | ACOTOCTICCA       | TOOTHOODO                       | ACCTOGOGY           |
|       | CACCTATOCO   | _                      | COMMANDE    | CCCCCCCCCCCC   | CCTCCCCCCCC        | CACTGCCCCT   | מכככתיומכיום      | TCCACCACCT        | ACCANCCCC                       | TOCAGCGG            |
| 11561 | Car and Carrier  | _                      | coractrico  | COCTOCTCCT   | CTTCCCGACT         | GRECATTICE   | TICTOTIANA        | GOCHDANANA        | DATCATOGRO                      | TCAOTCOM:A          |
| 10659 | COTOCCCAG  | _                      | _           | _  | GAAGGCTGA          | CCCCTAAAGG   | AAGAGGATAT        | cconcramin        | CTAGTACCTC                      | AGTCAGCTCF          |
| 21601 | ACARCIACAC   |                        | CCCTCTGAGT  | TEGECACEAE   | CONTRACACC         | GATGCCGCCA   | ACCCCCCTAC        | CACCTTCCCC        | OPCORDOCAC                      | CCCCGCTTGA          |
| 4     |  |                        | GGGAGACTCA  | -  | GCGCACACTTAG       | CTACGGCGGT   | TOCOCCENTO        | GTOGAAGGGG        | CAGCTCCCTO                      | COCCCCANCT          |
| 10110 | ACTION OF THE PERSON OF THE PE | _                      |             | _  | AGCGAAGACG         | ACCARGACCG   | CTCAGTACCA        | ACAGAGGATA        | AAAAGCAAGA                      | CCAGGGACAN'         |
| 10.69 | CONCERCE   |                        |             | TCCAMACAT  | TCCCTTCTGC         | TOCTOCTAGE   | GAGTCATOOF        | TOTOTOTAT         | THECHICA                        | CONCENCING          |
|       |  |                        |             |  |                    |              |                   |                   |                                 | -                   |
| •     |  |                        | Activities  | CHESTACGARA  | COCATOGORA         | CTACCTAGAT   | GTOCGAGACO        | ACGTOCTOTT        | BAAGCATCTG CAGCGCCAGT           | CAGCGCCAUT          |
| 73807 | CCACACACAM   |                        |             |  | CCOTACCGCT         | GATGGATCTA   | CACCCTCTOC        | TECACGACAA        | CTTCGTAGAC GTCGCGGTCA           | GTCGCCGTCA          |
|       | רפורוריפון   | -                      | -           | GCACCGATGE   | CECCTECATO         | ATAGCGGATT   | TCAGCCTTCC        | CTACGAACGC        |                                 | CACKTOCOCOT         |
| 70657 | CCCCALTAI  |                        |             | COTCOCTACA   | CYNOGAGORG         | TATCGCCTAC   | AGTOGGAACO        | GATOCITOCO        | GTCCATARCA                      | OTGGCGCGCA          |
|       |  |                        |             |  | COCCOCOLOR         | ACTICIACC    | CUTATTRICC        | <b>OTCCCAGAGG</b> | TOCTROCCAC                      | CTATCACATIC         |
| 24001 | ACCCCCCAM  | COCCAMBANA             |             | OCTCOORTIG   | COCCCCCACT         | TCAACATCAG   | CCATANACOO        | CACGOTOTOC        | ACGAACGOTO                      | GATACHCITAS         |
|       | Topologic I  |                        |             |  |                    |              |                   |                   |                                 | A CUBAN             |
|       |  | 44044                  |             | ACCOUNTS AND ACCOUNTS TO THE TRANSPORT ACCOUNTS ACCOUNTS | Acceesage          | ARCGGACANG   | CAGCTCCCCT        | TREFRECAGES       | CAGCTGGCCT TGCTGCAGG CGCTGTCATA | CCTCIATATO:         |
| 24101 | SABORGETT!   | TEACTIFCTA             | TOGOCATACG  | TOBOGATAGO ACGOCACGAT TOACGATCAGC                        | TGGCGTCGGC         | resectionic  | GTCGACTGGA        | ACCICCOTICCC      | ACRECETICE GEGACAGIAT           | CCALTATARY.         |

Figure 150

# PMRKAdSgag MERGRZ

| 10070                                   |              | COBACTICAL                | NAARTY TERE            | Acasemental  | ACGCCIACGAG  | ANGRESIA            | CAAACCCTCT                         | CCANCAGGAA    | AACAGCGAAA        | ATGRANCITCA     |
|---|--------------|---------------------------|------------------------|--------------|--------------|---------------------|------------------------------------|---------------|-------------------|-----------------|
| 10242                                   | TANA TRACTE  |                           |                        |              |              |                     |                                    | contorectt    | Trotcoctit        | TACTIFICAGT     |
|   |              |                           | XDeol                  |              |              |                     |                                    |               |                   |                 |
| 24101                                   | CHCTCACAGTO  | THEOTOGAGE TECHNOLOGICADA | TCGACGGTGA             | CAACIACATA   | CTAGCCCTAC   | TANAACTICAG         | CATCSARGETC                        | ACCCACTITO    | CCTACCCGGC        | ACTIAACCTA      |
|   | GAGACCICAC   |                           | AGCTCCCACT             |              | GATCCCATC    |                     | GTARCTCCAG                         | TYXIGHOAAAC   | GCATGGGCCC        | TRANTTEGAT      |
| 24401                                   | CCCCCCAAGG   |                           | AGTCATGAGT             | GARCTGATCG   |              | מנשטרנינדפ          | GAGAGGGATO                         | CAAATTTGCA    | NONACAARCA        | מאמפאמפני.      |
|   | GGGGGGTTCC   | AGTACTCGTG                | TCAGTACTCA             | CTCCACTARC   | ACCARTACACG  | CONTRACTOR          | CTCTCCCTAC                         | GFTFAMACGT    | TCTIGITION        | CHACH CLIM      |
| 24501                                   | TACCCOCAGT   | POCCARCAN                 | CAGCTAGGG              | מכונסטכידוכה | AACGCGCGAG   | CCTGCCGACT          | TODAMONDCO                         | ACCCAAACTA    | ATCATCOCCO        | CAGAICTC        |
|   | ATOOOCOTCA   | ACCECTOCIC                | GTCGATCGCG             | CCACCGAAGT   | TTK:COCCTC   | CONCONCITGA         | ACCTCCTCCC                         | TOCOTITIOAT   | TACTACCOOC        | GECACGAGEA      |
|   |              | 6                         | 7                      |              |              |                     |                                    |               |                   | • !             |
| 24601                                   | TACCOTOGAG   |                           | CITICAGISCA TGCAGCGGTT | CTTTGCTGAC   | CCGGAGATTGC  | ACCCATANGET         | AGNOTIAAAGA                        |               | CCTITICGACA       | COCCTACGTA      |
| l<br>                                   | ATOCACCTC    |                           | ACOTCOCCAA             | CANACCACTG   | GGCCTCTACG   | TCGCGTTCGA          | Tercentar                          | AACOTGATGT    | CCANACCTOR        | CCCGATGCAT      |
|   |              | Bytti                     |                        |              |              |                     |                                    |               |                   |                 |
| 24701                                   | CACCAGGCCT   | 8                         | CAACGTGGAG             | CTCTGCAACC   | TRESTUTECTA  | CCPTOSOAATT         | TTCCACGAAA                         |               |                   | CFTCATTCCA      |
| •                                       | GCOOLCCOOA   |                           | GTTGCACCTC             |              | ACCAGAGGAT   | GONACCTTAA          | MCGTGCTTT                          | TOOCOGNACC    | COTTTTGCAC        | GANGTANGOT      |
|   |              | Asci                      | 3                      |              |              |                     |                                    |               |                   |                 |
| 24801                                   | COCTCAAGOG   | B                         | COCGACTACO             | TCCCCCACTC   | COTTTACTTA   | TTTCTATGCT          |                                    |               | <b>GOCOTTTOCK</b> | ACCAGNOCTY      |
|   | CHECKAGTTCCC |                           | <b>OCOCTIGATOC</b>     | ACCCCCTOAC   | GCAAATGAAT   | ANNGATACGA          | TOTOGRACCOT                        | CTCCCCCTAC    | CCOCMACCO         | TCGTCACGAA      |
|   | •            |                           | PSI                    |              |              |                     |                                    |               |                   |                 |
| 24901                                   | CCACCACTCC   | NCCTCNOO                  | ACCTOCAGA              | ACTRICTAMAG  | CANANCTTGA   |                     |                                    | AACGAGCGCT    | ccaraccac         | GCACCTG3(1)     |
|   | CCTCCTCACO   | -                         | TEGACOTETT             | TCACCATTTC   | GTTTTCAACT   | TCCTGGATAC          | CTOCCROAND                         | TIMETICACION  | OGCACCOOCG        | COLUMBACCO      |
| 10020                                   | CACAMINATION | Tryphysham.               | CCTICATIVAN            | ACCUTACAAC   | MAKENCHACC   | AGACTTCACC          | AUTICAAAGCA                        | TOTTOCAGAA    | CTTTAGGAAC        | THEN PECTA      |
| 10003                                   | CHOTAGEAAA   |                           | GGACCAATTT             | TOGGACGITG   | TCCCAGACGG   | TETENAGTOG          | TCAGITTICGE                        | ACAACGICTT    | GANATCCTTG        | AAATARGATK'     |
| 10130                                   |              |                           | CACCACTER T            | GTGCACTTCC   | TARCCIACTET  | OTTOCCCATTA         | ACTACCGCGA                         | ATOCCCTCCO    | CCGCTTTGGG        | GCCACTOCTA      |
| 10167                                   | TOGGGAGICE   |                           | COCTOCACCA             | CACGTGAAGG   | ATCCCTGAAA   | CACGGGTANT          | TCATGGCGCT                         | TACGGGAGGC    | OCCOMPACCC        | COCHUNCONT      |
|   | lied.        |                           |                        |              |              |                     |                                    |               |                   |                 |
| 25201                                   | THE PRICAGO  | CETTERIOR CTAGGGACT       | ACCITOCCTA             | CCACTCTVAC   | ATAATCGAAG   | ACCITIANCES         | ACGINIANCES INACOSICIA CINSAGISTIC | CTOSAGTOTC    | ACTORCOCTO        | CAACCTATC       |
| 4                                       | GENARIACETE  | DATEGOTICA                |                        | GOTCAGACTO   | TATTACCTTC   | TOCACTORIC          | TOCACTORCE ACTOCCAGAT              |               | TGACAGCGAC        | GTTGFATAL       |
|   |              |                           |                        |              |              | \$                  | - F                                | Pul)          |                   | !               |
| 10126                                   | Acceptable   | acreection?               | TTOCAATTCO             | CACCTCCTTA   | ACCIANAGICA  | AATTATCGT           | ACCTITICAGE                        |               | CICOCCIGAC        | GANAGICCT       |
|   |              |                           | _                      | GTCGACGAAF   | TOCTITICAGE  | TTAATAGCCA          | TOGARACTEG                         | ACCIFICCIONOS | GAGCGGACTG        | CHITANGE        |
| LOADL                                   |              | _                         |                        | TOTOGOCOTO   | COCTTACCTT   | COCAMATTE           | TACCTGAGGA                         | CTACCACGCC    | CACGAGATTA        | COLLICTACOA     |
| 10567                                   |              | _                         |                        | _            | CCGNATGGAA   | COSSITIVANC         | ATCACACTCCT                        | GATOCTOCOG    | <b>GTGCTCTAAT</b> | CCAAGATUT       |
| 2022                                    |              |                           | _                      | •            | CTRCATTACCC  | ANTICICACAT         | TCTTVARICAN                        | TTGCMGCCA     | TCAACAAAGC        | CCCCCMON        |
| *************************************** | TCTOOTTAGO   | _                         |                        | _            | CANTAATY     | <b>Trecessition</b> | ACAACCOGTT                         | AACOTTCGGT    | ACTIGITIES        | וש'ארנאיז זר זו |
| CARC                                    |              | CANADOMAD.                | CENTRAL                | THECACTOCC   | ACTROCOCATOR | GCACCTCAAC          | CCAATCCCCC                         | COCCOCCOCA    | OCCUTATION        | CAMERICA        |
| 43004                                   | AMGACGATG    |                           |                        | _            | TCAUCCCGCT   | CCTCGAGIFTG         | CCTTACKA                           | ACCIONACOT    | CCKCATAGTC        | Greeneraco      |

Figure 15F

#### PMRKACTSgag NER682

Psff

|       |                   |                   |                   | S.A  | . 4                   |                       |                   |                    |                       |                    |
|-------|-------------------|-------------------|-------------------|--|-----------------------|-----------------------|-------------------|--------------------|-----------------------|--------------------|
| 25701 | 200000000         | TRECCAGGAT        | <b>OCCACCTANA</b> |  | AAGAAGCTOC NATHALYAYC | וניגעיגנעים           | CACGAGGAGG        | AATACTOCOA         | CAMTCAGGCA            | CARRACACTETE       |
| •     | CCCOOGNACO        | AAGOGTCCTA        | CCURRETIFF        | TTCTTCGACG   | Terthanology          | CLYCTYNGOTYSC         | CTGCTCCTCC        | TTATOACCCT         | <b>STCACTCCOT</b>     | CTCCTCCN .         |
|       |                   |                   |                   |  |                       | Hariff                |                   |                    |                       |                    |
| 25801 | TOCACCIACIA       | COACCACCAC        | ATGATTGAAG        | ACTROCIAGAG  | נענעשעשעענע           | C                     | AGCITCISAGA       | CGTOTCACAC         | GAMACACCUT            | CACCOTACOV .       |
|       | ACCTGCTCCT        | cerecreero        | TACTACCTTC        | TGACCCTICTC  | ממארידמניזני          | CTTCGAAGOC            | TCCAGCTTCT        | CCACAGACAG         | CTTTGTGGCA            | מומטביים מו        |
| 25901 | CGCATTCCCC        | Traccoacac        | CCCAGANATC        | GGCAACCGGT   | TOCACICATION          | CTACAACCTC            | CACTECTERAG       | GCGCCGCCCG         | CACTGCCCOT            | TOTACODACCC        |
|       | OCOTANGOCO        | AGCGGCCCCC        | GOTTETTAG         | CCCATTORCCA  | AGGIRGOTACC           | CATESTAGO             | CCGAGGAGTC        | cacacacac          | GTCACGOGCA            | ADCCCCTCCO         |
| 26001 | AACCOTAGAT        | GOODCACCAC        | TOOMCCARR         | OCCUGINAGE   | CCAMICAGC             | CCCCCCCTTA            | CCCCANGAGC        | AACAACAGCG         | CCAAGOCTAC            | CINCINCATION"      |
|       | TIOCCATCTA        | CCCTCTCGTG        | ACCITICGUCC       | CONCATTON  | GGTTCGTCGG            | CTCCTTCANT            | CCCCTTCTCC        | THETHOREGE         | GGTTCCGATG            | <b>OCCIMPTACCU</b> |
| 26101 | <b>OCCOOCACAA</b> | GAACGCCATA        | OTTOCTTOCT        | TOCANOACTO   | TOCORCICANC           | ATCTCCTTCG            | CCCGCCGCTT        | TUTTUTCTAC         | CATCACGUCG            | TOXIA THUCK        |
|       | COCCCOION         | CTTCCCCTAT        | CHACCHACGA        | ACCITICAGAC  | ACCCCCGTTG            | TAGAGGAAGC            | GCCCCCCCAA        | AGAMINGATO         | GTAGTGCCGC            | ACCHINARION.       |
| 26201 | CCOTAACATC        | -                 | ACCONCATCT        | CTACAGCCCA   | TACTRICACER           | CCCCCACAGCCG          | CAGGNACAGG        | ACCOCCACA          | CAGAAGCAAA            | CHACCIACCOCA       |
|       | OCCATIGING        | GACOTANTGA        | TOOCAGTAGA        | CATCTCOCCT   | ATGACGTYCC            | כואככבורפככ           | GICGINGICG        | TCCCCGGTGT         | <b>greatents</b>      | CCGCTGGCC1         |
| 26301 | TAGCANGACT        | CTGACAAAGC        | CCAAGAAATC        | CACAGCGCG  | <b>GCARCARCAR</b>     | GACCIACTACC           | <b>octoconcra</b> | GCGCCCAACO         | AACCCOTATC            | GACCCOCOAR         |
|       | ARCOFICIOA        | <b>GACTOTTICO</b> | <b>GGTTCTTTAG</b> | GTGTCGCCCC   | corcorcorc            | CTCCTCCTCG            | CONCOCAGAC        | cacacamac          | TTOOOCATAG            | CTOGGCGCTC         |
| 26401 | CTTAGAACA         | GUATTETECC        | CACTOTOTAT        | CCTATATIFIC  | AACAGAGCAG            | GROCCANDA             | CANGAGCTGA        | AAATAAAAA          | CADOTETETO            | COATCCCTCA         |
|       | GAATCTFTGT        | CCTANAAAGG        | GTGAGACATA        | CCATATAAAG   | TRACTORIC             | CCCGGTTCTT            | GTTCTCGACT        | TITATIFIE          | GTCCAGAGAC            | <b>OCTACOGART</b>  |
| 26501 | CCCCCAGCTO        | CCTOTATCAC        | AAAAGCGAAG        | ATCAGCTTCG   | GCGCACGCTG            | האחחהפכפס             | ARACTECTET        | CACTAAATAC         | TOCCCCACTOR           | CTCTTAAGGA         |
|       | GOOCGTCGAC        | CCACATACTO        | Trincocrac        | TAGTCCAAGC   | CCCOMOCIAC            | CTTCTGCGCC            | TECGAGAGAA        | <b>OTCATITIATO</b> | ACCCCCCACT            | GAUAATTEC .        |
| 26601 | CHAOTIFICOC       | OCCUPANT          | AAATTTAAGC        | OCCUMANCTA   | COTCATCTCC            | ACICOCICCACA          | CCCGGCGCCA        | OCACCTOTTO         | TCAGCGCCAT            | TATHINGCANG        |
|       | GATCANAGEG        | COCCANACAC        | TITAAATICG        | CCCTTTTGAT   | CCACTACAGG            | reaceasar             | OPOCEGEOUT        | CCTGGACAAC         | Acresector            | ATACTICGTTC        |
| 26701 | GAMATICCCA        | COCCCTACAT        | GTCGAGTTAC        | CARCCACAAA   | TOGGACTICA            | GOCTGGAGCT            | <b>OCCCANGACT</b> | ACTICAACCCO        | AATAAACTAC            | ATTOMICOCON        |
|       | CTTTAAGGGT        | OCCOONTICTA       | CACCTCAATO        | GICCONGITT   | ACCCTGAACO            | CCGACCTCGA            | COCOTTCTGA        | TCACTTGGGC         | Tratticato            | TACTCGCGCC         |
|       |                   | Ecoffy            |                   |  | i i                   | ILO2                  |                   |                    |                       |                    |
| 26801 |                   | GATATCCCGG        | OTCAACGOAA        | TACOCOCCCA   | CCCIAAACCCA           | CCCAAACCCA ATTCTCCTCG | AACAGGCGGC        | TATTACCACC         | ACACCTCOTA            | ATAACCITIAA        |
|       | CTOCOCOTOTA       | CTATAGAGE         | CAGITICALT        | ATGCGCGGGT   | ואלודואלו             | TANGAGGACC            | THETECOCCE        | ATANTOOLOG         | ומומשתאו              | TATACANTT          |
| 26901 | <b>TCCCCOTAGT</b> | TOOCCCOCTG        | CCCTOGTGTA        | CCACCOLAROT  | ככבטכעכעכע            | CCACTGTGGT            | ACTTCCCAGA        | _                  | CCGAAGTICA            | GATTACTAN!         |
|       | ACCOCCATICA       | ACCOGGCGAC        | GGGACCACAT        | <b>GOTCCTFTICA</b>   | GOCCACOCT             | GUITGACACCA           | TCAACCCTCT        | CIRCOGOSTICC       | GOCTTICANGT           | CTACTCATTO         |
| 27001 | PCADODOCOC        | AGCTTGCGGG        | COCCITICOT        | CACAGOOTIC   | ממזכמככנה             | CCAGGGTATA            | ACTICACCTICA      | CANTCAGAGG         | <b>OCCADOTATE</b>     | CARCTCAACT         |
|       | AGTECECOES        | TCGAACGCCC        | GCCGMANGCA        | OTOTOCCACO   | CCARCAGACC            | CCTCCCATAT            | TOAGTGGACT        | Ξ.                 |                       | OFCONGTINE.        |
| 27101 | ACGAOTCOOP        | GAGCTCCTCG        | CHOOPETEE         | GICCOGACGO   | GACATTICAG            | ATCGGCGGCG            | ccaaccactc        |                    |                       | CANTECTANE         |
|       | TECTCAGCCA        | CTCGAGGAGC        | DAACCAGAGG        | CAGGCCTGCC   | CTCTAAAGTC            | TAGCCCACCCAC          | GGCCGGCGAG        | AAGTAAGTGC         | GGAGCAGTCC            | GTTAGGATTG         |
|       |                   |                   |                   |  |                       |                       |                   | _                  | A Comments & Comments |                    |
| 102/2 | ACACCAGACC        | AGENCETCHE        | ACA:COCGCTC       | ACCTICATION AND ACCTICATION AC | CCTTCAGACT            | TTANTANCT             | CCTCANACAC        | GGTMGCCAGA .       |                       | GWGAGGCCT          |
|       | 1                 |                   |                   |  |                       |                       |                   |                    |                       |                    |

Figure 1502

| 101.66      |                  | ACTATOCOM      | TCAATTTATT  | CCTANCTING        | Archanta                       | ההאכינים נ                      |                        |              | AAGTOOAGAG   | GCAGAGCAAC    |
|-------------|------------------|----------------|-------------|-------------------|--------------------------------|---------------------------------|------------------------|--------------|--------------|---------------|
|             | CONTROCCO        | TGATAGGCCT     | AGTTANATAN  | CHATTONANC        | אניניגינ:אדוד<br>אניניגיניאדוד |                                 | -                      | TOACTTACAA   | TICACCICIC   |               |
| 27401       | TOCOCCTGAA       | ACACCTGOTC     | CACTOTORCAC | CCCACAACTG        | GANALGARA                      | האנידור כמסדם י<br>כידמאה מככחב | ACTITIVACTA TENNANCENT | GAACTTANG    | OGGCTCCTAG   | TATACCTC(1    |
| 27501       | CCCGGCGCAC       | -              | TACCGCCCA   | COCAGACCTT        | מכככנינוענייכ                  |                                 |                        |              | TACTFORGED   | CICACACICO :A |
|             | <b>COCCCCCTO</b> | CCCCAGGCCG     | AATGGCCGGT  | CCCTCTCGAA        | CCCCCATCGG                     | ננט                             | CAAATGGGTC             | CCGGGGGACG   | ATCARCICOL   |               |
|             |                  |                |             |                   |                                |                                 |                        |              |              |               |
| 27601       | ccctororic       | TCACTOTOAT     | TTGCAACTOT  | CCTAACCCTG        | GATTACATCA                     |                                 |                        | GIGCITAGIA   | TANTAAATAC   | _             |
| 1           | GOODCACANG       | A AGTGACACTA   | AACCITICACA | GOATTROGRAC       | CTAATGTAGT                     |                                 |                        | CACGACICAT   | ALTAITINIO   |               |
| 27701       | ATATACTOOD       | OCTCCTATED     | CCATCCTGTA  | AACCCCACCG        | Trittacces                     |                                 |                        | CCTTACCTOO   | TACTITIONS   | MERCICO       |
| •           | TATATOACCC       | CCAGCATAGC     | GGTAGGACAT  | TIGGGGTGGC        | AGANGTGGGC                     | _                               |                        | CCAATCCMCC   | ATGAGATTO    | -             |
| 27801       | CHOTOATTTA       | CAACAGTTIC     | AACCCAGACG  | CARTOAGTET        | ACCINCIACIANC                  | CHCHCCCMOC                      |                        | CATCAGAAAA   | AACACCACCC   | TCCTTACCT     |
| )<br>)      | GACACTAMAT       | -              | TOCOTOTO    | CTCACTCAGA        | TIGGRETICATION                 | GAGAGGCTCO                      |                        | GFAGICTIT    | oonioniou.   |               |
| 27901       | CCOCONCOT        | ACGNOTOCOT     | CACCGGCCGGC | TOCACCACAC        | CTACCGCCTG                     | ACCOUNTACC                      |                        | COCACAGACC   | TCANTAACTC   | ACABATTETE A  |
| •           | COCCCTTOCA       |                | Gradecode   | ACOTOGRAND        | CATOCCOGAC                     | TOCCATTING                      | TCTGANAAAG             | BCCTCTCTV6   | Well Allow   |               |
| 10082       | ARCAGGGGGT       | GAGCTTAGAA     | MCCCTTAGG   | <b>OTATTAGGCC</b> | ANARGEGERAG                    |                                 | GTTTATGAAC             | ANTHCANOCA   | ACTICTACOGO  | CTATTCTAAT    |
|             | TIGHTCTICCA      |                | TTOOGNATCC  | CATAATCCOO        | Trrececore                     | GATCACACCC                      | CAANTINCTTO            | TRACTICOT    | TOMONTOCCC   |               |
|             |                  | Xtbal          |             |                   |                                |                                 | •                      |              |              |               |
| 28101       | TCAGGITTET       | CTACAATEGO     | COMOGGG     | ATTCTCTGTC        | TRETGATTET                     |                                 | ATACTAACOC             | TICHCIGCCT   | ANGOCTCOCC   | OCCUPACION :  |
|             | ACTECANAGA       |                | CCAACCCCAA  | TAAGAGACAG        | AACACTANGA                     | GANATARGAA                      | TATEATTEC              | MAGAGACCOA   | TICCARCCO    |               |
| 28201       | Trac ACATETRO    |                | CACCTITITA  | AACGCTGGGG        | TCGCCACCCA                     |                                 |                        | CTAGGTTTAC   | TCACCCTTGC   | GICAGUCCAU    |
|             | ACOTOTANAC       |                | GTCGAAAAAT  | TIGCGACCCC        | AGCGGTGCGT                     | TCTACTAATC                      | CATCTATTAG             | CATCCANATO   | ACTICOLOGICA |               |
|             | Kont             |                |             |                   |                                |                                 |                        |              | PETERARTY.   | ACCACAGA!     |
| 28301       | COTACCACC        | S ANAGOTOGA    |             |                   |                                | COCAGCTGAA                      |                        | CACCACIC!    | ATATITACO    |               |
| •           | CCATGGTGGG       | 3 TTITICCACCT  | AAAATTCCTC  | <b>GOTCCCACAT</b> | TACAATGTAA                     | GCCTTCGACTT                     |                        |              | Chemina      |               |
| 28401       | ATGAAAAGCT       | r octriviticae | CACAAAAACA  | MATTOCCAA         | GTATCETCIT                     | TATOCIATIT                      | <b>MICAGCCAOG</b>      | TCACACTACA   |              |               |
| 1<br>3<br>9 | TACTITITICAL     |                | GIGHTTIME   | TTANCCOTT         | CNTACGACAA                     | ATACCATANA                      | CCOLCOGUCG             | ACTION OF    |              |               |
|             |                  |                | E G         | Bs111071          |                                |                                 | !                      |              |              |               |
| 28501       | CEAGGGTAA        | A ACTICATABAA  | CITITATOTA  | TACTITICCA        | TITTATEAM                      | TOTOCCACAT                      |                        | ATGAGGAGAG   |              |               |
| 4000        | CONCECATIT       |                |             | ATCAMAGGT         | ANNTACTIT                      |                                 | _                      | TACTOUTUR    |              |               |
| 28601       | CAAAATTGTG       | -              | TOTAL       | -                 |                                |                                 | CETTROPICE             | GTACCCTACT   | CTATATTAGA   |               |
|             | CTTTTACAC        | _              | ACCOTONAG   | ACGACGTGAC        | CATACCATTA                     | -                               | CCAMICLAGA             | CATGOOMIGA   |              |               |
| 28701       | GACGCAGCTT       | -              |             | -                 |                                |                                 | ACCACTAACT             | OCT TRACTOR  | GACCAROCAN   |               |
|             | CTOCOTOCAL       | A ATAACTCCTT   | TACTITIAGG  | -                 |                                | PCGATTACM3                      | אניו ואייאייאי         | Para and the |              | -             |
| 28801       |                  | _              | -           |                   | OCHCATTICC<br>CCACCANANCO      | ACCACTCANTAC                    | GTAAGGGGAC             | TTOTTANCTO   |              | -             |
|             | THICANCO         | O TANTATTAAT   | CHANCCIAN   | ATTICOGGG         |                                |                                 |                        |              |              |               |

Figure ISR

## PHRKAdSgag MER682

| . 20901 | CCCCTACARC         |                       |             | CTTGAAGTCA GGCTTCCTGG ATGTCAGGAT CTRAACTTTKA CCAGGACCTG | וויישני וויישנייי | CCAGCACCTO                        |                        | PCCCACCACAT FIGHTCCAGT  | CONCINCION                              | CCACCCACTC         |
|---------|--------------------|-----------------------|-------------|---|-------------------|-----------------------------------|------------------------|-------------------------|---|--------------------|
|         | COCCATOTTO         |                       | CCCMICCALC  | TACAGTECTA  | CACTGAAACC        | CONTECTOR                         | AGGGGGCCTA             |                         | ANCANGETCA GETTERATERY                  |                    |
| 29001   | TANCAGAGAT         | GACCAACACA            | ACCAACIGGG  | בבפכבמניבעב   | CCRONCTIACA       | TUTACCACA                         | ATACACCCCA             | ACTITICACC              | TTTOTCAATA                              | ACTOGGNIAA         |
|         | ATTOTOTOTA         | CREGITOTOF            | TOOTTOCCC   |   |                   | AGATIATICITY                      | TATIONSCOOT            | TCAMOACCC               | AAACAGTTAT                              | TOACCCTAIT         |
| 29101   | CTTCCCCATC         | TOOTOGFTCF            | CCATAGCGCT  | TATISTA   | TOUTTAIN          | TTATGIVECT                        | CATCTGCTGC             | CTARACCCCA              | AACOCGCCC                               | ACCACCCATC         |
|         | <b>GAACCCOTAC</b>  | ACCACCAAGA            | OCTATOGOGA  | ATACAAACAT  | ACCOMMENAT        | ANTACACYOA                        | GTAGACGACG             | GATTECOCOT              | MCCGCCCGC                               | TOCTIVETING        |
| 29201   | TATAGECCA          | TCATIOTCCT            | ACACCCAAAC  | AATTATTATA  | TCCATAGATT        | CATACCACACTO                      | ANACACATTE             | TUTTICIET               | TACACTATOA                              | TTANATTCACIA       |
|         | ATATCAGGGT         | AGTAACACGA            | TOTOGGITTO  | TTACTACCIT  | AGGTATCTAA        | CCTGCCTCIAC                       | THEFT                  | ACAAAAGAGA              | ATCTCATACT                              | AATTTACTCT         |
|         | ţ                  | Xhci                  |             |   |                   |                                   |                        |                         |   |                    |
| 29301   | CATCATTCCT         | CATCATTCCT CCACTTTTTA | TATTACTOAC  | CCFIGNICE   | CTTTTTTG          | CTITITIONS CONSCINE               | ATTORCTOCO             | OTTICICACA              | TCOMOTAGA                               | CTCCATTC A         |
|         | . GTACTAAGGA       | OCTCAMANT             | ATAATGACTO  | CCAACAACCC  | CAAAAAAACAC       | CAAAAAACAC GCACGAGGTG             | TAACCGACGC             | CAAAGAGTOT              | AGCTITCATES                             | GACOTANO .T.       |
|         |                    |                       |             |   | P <sub>S</sub> () | 10                                |                        |                         |   |                    |
| 29401   | <b>OCCTITCACAG</b> | <b>1CTATTTGCT</b>     | TTACGGATT   | DICACCCTCA  | CGCTCATCTO        | CRETEATETR CAGGETEATE             | ACTIGICATICA           | TCCCCTTTAT              | CCARTGCATT                              | GACTOGOTHE         |
|         | COGARGTOTC         | ACATAAACGA            | AATOCCTAN   | CACTICACAGT   | GCGAGTAGAC        | GTCGGAGTAG                        | TGACACCAGT             | TCACACCAGT AGCGGAAATA   | GOTCACGTAA                              | CTGACCCANA         |
|         | •                  | ٠                     |             |   |                   |                                   | E:co                   | 22.0                    |   |                    |
| 29501   | CHOROCOCH          | TOCATATOR             | AGACACCATC  | CCCAGTACAG  | GGACAGGACT        | ATAGCTOAGC                        | THETTAGANT             | THETHAGANT TETTTAATTA   | TOWATTERC                               | TOTORCETT:         |
|         | CACACGCGAA         | ACCTATAGAG            | Teterrated  | <b>GOOTCATISTS</b>                                      | CCTGTCCTGA        | TATEGACTEG                        | ANGNATICITA            | AGAAATTAAT              | ACTITITABATO                            | <b>ACACTICAAA.</b> |
| 29601   | CTCCTCATTA         | THYOCACCCT            | ATCTGCGTTT  | TOTTCCCCGA  | CCTCCAAGCC        | TCAAAGACAT                        | ATATCATGCA             | GATTICACTEG             | TATATOGAAT                              | ATTCCAAGI T        |
|         | DACCACTANT         | AAACOTOGGA            | TAGACGCANA  | ACANDOCCCT  | CGACCTTCUC        | AGTITICIOTA                       | TATAGTACGT             | CTAAGTGAGC              | ATATACCTTA                              | TAAGGITCAA         |
|         |                    |                       | ٠           |   |                   |                                   | Pall                   |                         |   |                    |
| 29701   | OCTACAATGA         | AAAAAGCGAT            | CTTTCCGAAG  | CCTCGTTATA  | TCICAATCATC       | TCTCTTATAG                        | TOPPICAGO              | TOTTICTISCAS TACCATCTTA | OCCUTAGOTA                              | TATATCCCIA         |
|         | COATOTTACE         | TTTTCCCTA             | GAAAGGCTTC  | GOACCAATAT  | ACCTTAGTAG        | AGACAATACC                        | ACAMGACGTC             | ATOGTAGAAT              | COCCATCOAT                              | ATATAGGAT          |
| 29801   | CCTTOACATT         | OCCHOOLACE            | CAATRGATGC  | CATGAACCAC  | CCANCTITICC       | כבפבטבבבופב                       | TATOCTTCCA             | CTGCAACAAG              | TTOTTOCCOG                              | COCCITITION        |
|         | DOMCTOTIVA         |                       |             | GTACTTGGTG  | COTTICALANCE      | COCCCCCCCC                        | ATACGAAGGT             | GACGITOTIC              | AACAACGOCC                              | GCCCANALA 1        |
|         |                    | ٠                     |             |   |                   |                                   |                        |                         | × * * * * * * * * * * * * * * * * * * * | );if               |
|         |                    |                       |             |   |                   |                                   |                        |                         | 1004                                    |                    |
| 29901   | CCAGCCANTC         | Agenteacte            | ACCTIONCE   | ACCCCCACTO  |                   | CTTTAATCTA                        | ACAMANAGAO, ATGACTGACA | ATGACTOACA              | CCCTAGATCT                              | AGAAATTGCAC        |
|         | CONCOUNTAG         | TCGGAGCGG             | TOGMICACIC  | TERCOGNANC  | TTTAGTCGAT        | GANATTAGAT                        | reference Tactgacter   | TACTGACTGT              | GOCATICTACA                             | terrracero         |
| 30001   | GGAATTAITA         | CAGAGCAGCG            | CCTCCTAGAA  | AGACGCACAG  | CAGCOGCCGA        | CCANCAGCGC                        | ATGAATCAAG             | AGCTCCANGA              | CATOOTTAAC                              | TTGCACCAGE         |
|         | CCTTAATAAT         | grenconcec            | GCACGATCTT  | TCTOCOTCCC  | OTCOCCOCT         | COTTOTOGO                         | TACTINGTIC             | regraphics              | GTACCAATTG                              | AACTHOOTICA        |
| 30101   | GCANANOOO          | PATCHINGS             | CTCGTAMAGG  | AGGCCANAGE  | CACCTACGAC        | AGTAATACCA                        | CCGGACACCG             | CCTTAGCTAC              | AAGTTGCCAA                              | CCANGCOT!"         |
|         | COTTITICCCC        |                       | GAGCATTTCO  | TCCCOTTTCA  | CTCCATCCTC        | TCATTARKET                        | <b>GCCCTGTGGC</b>      | GONTCOATO               | TTCAACOOTT                              | GETTICITAL:T       |
| 30201   | CONNTROCTO         | OFCATOOTOG            | CACIANAACCC | CATTACCATA  | ACTEAGEART        | CHATTACHANC                       | CANAGOCTOC             |                         | CTTOTCAAGO                              | ACCTOARGAT         |
|         | CTTTAACCAC         | CAGTACCACC            | CTCTTTTCGG  | <b>CTANTISCTAT</b>                                      | TGAGTCGTGA        | <b>OCCATUTION</b>                 | GCTTCCGACG             | TAAGTGAGTG              | DAACAGTTCC                              | TOGACTOCTA         |
|         |                    |                       |             | Byff  |                   |                                   |                        |                         |   |                    |
| 30301   | CTCTOCACCC         | TTATTANGAC            | CCTOTOCOOT  | 3   | FTATTCCCTF        | TAACTINATAA AAAAAAATAA TAAAGCATCA | ANAMANATA              | THANGCHTCA              | CTTACTTAAA                              |                    |
|         | GAGACOTOGG         | AATAATTCTG            | GGACACGCCA  | GGACACGCCA GAGTITCTAS ANTANGGRAM                        |                   | ATTGATTATT                        | ተናንተባተለፕፕ              | ATTITOGINGI             | CANTGANTIT                              | TAGTCAATCG         |

# рикклабану непбв2

| 30401       | AMATTHCTUT        | CCAGTETATE  | CAGCAGCACC        | rectrosect   | CCTCCCAGGT         | CHATTATION        | AGCTTCCTCC   | TOCCTOCALA          | CHITCTCCAC        | ANTCTAAATG    |
|-------------|-------------------|-------------|-------------------|--------------|--------------------|-------------------|--------------|---------------------|-------------------|---------------|
|             | TTTANACACA        | COTCANATA   |                   | ACCAACCCCA   | CHIACKWITCCA       | CACTATAACO        | TYCANGGAGG   | ACCOACOTIT          | CHANGACOTO        | TTAGATTTAC    |
| 30501       | CONTRICACT        | TECTOCION   | TCCTGTCCAT        | CCCACACCCAC  | TATCTTCATG         | THEFTERTAGA       | TGAAGCGCGC   | AAAACCGTCF          | CHACATACCT        | TCAACCCCC     |
|             | CITACAGICA        | ANCIONODACA | ACCACACCTA        | CARCETORICE  | ATAGAAGTAC         | <b>ANCANOTHET</b> | ACTIVE TO CO | TTCTOGCAGA          | CFFCTATOGA        | ACTITIONS A   |
| 30601       | GTATCCATAT        | GACACOCAAA  | CCGGTCCTCC        | AACTAGTAGCCT | THETTACTE          | CICCCTITION       | ATECECCAAT   | <b>OCCUPITICANO</b> | AGAGITCCCCC       | TOCCOTACT:    |
|             | CATAGGTATA        | characentr  | <b>GGCCAGGAGG</b> | THGACACGGA   | ANGANTONG          | GAGGGAAACA        | TAGGGGGTTA   | CCCAAAGTTC          | TETEAGGGG         | ACCCCATGAG    |
|             |                   |             |                   | Ÿ.           | the<br>nume        |                   |              |                     |                   |               |
| 30701       | Termococc         | TATECGAACE  | TCTANTTACC        | TCCAATCCCA   | THICHTACKET        | CAMMATCACC        | AACTOTOTOT   | CTCTGGACGA          | OCCOCCANC         | CTTACCTCCY:   |
| ı           | ADAMCGCGG         | ATAGCCTTOG  | AGATCAATGO        | ACCTTACCGT   | AFGAACGCGA         | GTITACCCG         | TICCCCGAGA   | GAGACCTGCT          | CCGGCCGFTG        | GANTOCAGE 1   |
| 30801       | AMATOTAAC         | CACTOTORGE  | CCACCTCTCA        | AAAAAACCAA   | <b>GTY NANYATA</b> | AACCTOGAAA        | TATCTOCACC   | CCTCACAGIT          | ACCTCAGAAG        | CCCTANCTART   |
|             | TTTTACATTO        | GTGACACTCG  | COTOCACACT        | TITITION     | CAGTITICIAL        | TTCKIACCTTT       | ATAGACCIOS   | CCAGTGTCAA          | TOGRAFICTIC       | GOONTIGACA    |
| 30901       | <b>GOCTOCODOC</b> | GCACCTCTAA  | TOGTEGGGG         | CAACACACTC   | ACCATOCAAT         | CACAROCCCC        | CCTARCCGTG   | CACGACTCCA          | AACTTAGGAT        | TOCCACCCAA    |
| )<br>)<br>) | CCGACGGCGG        | -           | ACCAGCGCCC        | GTTCTCTCAG   | TOSTACCTTA         | GTOTCCGGTG        | CCATTOCCAC   | GTGCTGAGGT          | TTGANTCGTA        | ACCONTOCAL I' |
| 31001       | GGACCCCTCA        | CAGTOTCAGA  | AGGNANOCTA        | OCCCTOCAAA   | CATCAGGCCC         | CCTCACCACC        | ACCGATAGCA   | GTACCCTTAC          | TATCACTOCC        | TCACCCCCTT    |
|             | CCTOGGGAOT        | GICACAGACT  | TECTTTEGAT        | COCCACCTIT   | GTACTCCCCC         | CCACTCCTCC        | TOCCTATOOT   | CATOCOANTO          | ATACTOACOG        | AGTOCOCOA     |
| 31101       | TAACTACTOC        | CACTOGTAGE  | TTGGCCATTG        | ACTTORANGA   | <b>OCCCATITAT</b>  | ACACAMANTO        | GAMACTAGG    | ACTAANOTAC          | OCCCUCAT          | TOCHEGIAN .   |
|             | ATTOATOACO        | GTOACCATCO  | ACCCCTARC         | TOANCIFICT   | COCCETANATA        | TOTOTITIAL        | CTTTTGATCC   | TOATTICATO          | CCCCGNOGAA        | ACCTACATIVI   |
| 31201       | AGACGACCTA        | AACACTITGA  | CCGTAGCAAC        | TOOTCCAGOT   | <b>OTCACTATTA</b>  | ATAATACTTC        | CTTGCMACT    | AMOTTACTO           | GAGCCTTOGG        | THIRDATICA    |
|             | TCTOCTODAT        | TIGHTOMACT  | OCCATCOTTG        | ACCAGGTCCA   | CACTGATAAT         | TATTATONAO        | GAACGITTICA  | TTTCAMTOAC          | CTCCCCAACCC       | NANCTAN .     |
| 31301       | CAUCOCAATA        | TOCANCITAN  | TOTACCAGGA        | CONCTANCA    | TRATTCTCA          | NACAGACCC         | CTTATACTTO   | ATOTTACTTA          | <b>fccorradat</b> | GCTCANAAC:    |
| 1           | GTTCCOTTAT        | _           | ACATCGTCCT        | CCTGATTCCT   | AACTAAGAGT         | rmoreteed         | GAATATGAAC   | TACANTCANT          | ACCCARACTA        | CONOTITION    |
| 31401       | AACTAAATCT        | ANGACTAGGA  | CAGGGCCCTC        | TITITATANA   | CTCAGCCCAC         | AACTTOCATA        | TTAACTACAA   | CANAGOCETT          | TACTIONITY        | CAGCTTCAA     |
| ,<br>,      | TEATTTAGA         | ٠.          | GTCCCGGGAG        | ANANTATT     | averedecte         | TTGAACCTAT        | AATTGATGTT   | OFFICCOOAA          | ATGAACAAT         | GICCOARGET (  |
|             |                   | Hendill     |                   |              |                    |                   |              |                     |                   |               |
| 31501       | CAATTCCAAA        | AACCTTGAGG  | TTAACCTAAG        | CACTGCCANG   | CONTROATER         | TTCACCCTAC        | ACCCATAGCC   | ATTAATGCAG          | GAGATOGGCT        | TOWERTON      |
|             | OTTANGETTE        | -           | AATTOGATTC        | GIGACGGITC   | CCCMCTACA          | AACTOCGATG        | receratedo   | TAATTACGTC          | CICTACCCGA        | ACTIVAACCA    |
| 31601       | TCACCTAATO        | CACCAMCAC   | AAATCCCCTC        | AAAACINAAA   | THYRICCATOR        | CCTAGAATTF        | GNTTCAAACA   | ACCCTATCOT          | TCCTANACTA        | CONTINUE      |
|             | ACTOGATTAC        | _           | TTTACCCCAG        | TITIOTITI    | AACCEGIACC         | GGATCTTAAA        | CTAAGMTTGT   | TCCGATACCA          | ACCATITICAT       | CCTTGACCTR3   |
| 31701       | Tracerroy         | CACCACACOOF | GCCATTACAG        | TAGGNANCA    | ANTANTOAT          | MGCTAACTT         | TYTERROCAL   | ACCAGCTCCA          | TCTCCTAACT        | GTAGACTNAA    |
|             | _                 | _           | COGTANTOTIC       | Ancernett    | TTATTACTÁ          | TICGATTGM         | ACACCTRAFTG  | TOCTCOAGGT          | ACACCATICA        | CATCTGATTT    |
| 11801       |                   |             | TCACTITION        | CTTANCALA    | TOTAX              | AAATACTTGC        | TACACITITICA | OTTINOPERO          | TTAAAGGCAG        | TTOGCTCCA     |
|             |                   | : -         | AGTCAAACCA        | GAATTGTTTT   | ACACCUTICAG        | TITATICAACG       | ATCTCANGE    | CAAAACCGAC          | AATTICCOTC        | AAACCGAGGT    |
| 11.001      | APPRICATION.      | CACHTICAAAG | TESTERICITY       | ATTATAGAT    | TTCACCAAAA         | TERRETECTA        | CTANACANT    | CCFFCCFCCA          | CCCAGAATAT        | A. JANCITTA   |
|             | TATAGACCT         | GECALOTTE   | ACCINGTAGAA       | TAATATTETA   | ACTOCITIE          | ACCTCACGAT        | CATTTGTTAA   | GOLAGGACCT          | OCCUPATA OCCUPATA | ACCTTUMMT     |
|             | •                 | D. C.       |                   |              |                    |                   |              |                     | ,                 |               |
| 32001       | GUNATOGINA        | TETTACTORA  | OCCACACICE        | ATACAMACOC   | TUTTONIATE         |                   | TATCACCITA   | TCCAAAATCT          | CACCOTANAA        | CTGCCAAAAG    |
|             | CTTACCTCT         | AGNATICACTT | CCCTCTCTCGA       | PATOTTICE    | ACAACCTAAA         | TACCICATION       | ATAGTEGAAT   | ACCITITAGA          | GEGCCATETE        | GACCOSTITIC   |

Figure 15T

# PHRKAUSpag Merbiz

| 32101 | TANCATTOTIC  | AGTICANGITITE TCAGTTCAAA                             | ACTTANACCO<br>TGANTITYCC                             | ACACADANCT<br>TCTCTTTTTCA                      | AAACCITCITAA<br>TTTKX:ACAIT                          | CACTAMCCAT   | TACACTAAAC   | GCTACACAGG<br>CCATGTGTCC           | AAACAGGAGA  | CACAACTC!A<br>GTGTTRAG:"                                |
|-------|--|--|--|--|--|--|--|------------------------------------|---|---|
| 32201 | AGTGCATACT<br>TCACGTATGA                             | CTATOTCATT   | TTCATCCCAC   | Transcendent<br>Arcanacing                     | ACANI TACAT<br>TETTKIATTITA                          | TANTGAMATA<br>ATTACTTTAT   | PPICATCACAT<br>NANCCACHERA   | CCTCTTACAC                         | TETTICATAC<br>AAAAAGTATO                          | ATTGCCCAN:<br>TAACGCTTT                                 |
| 32301 | AATAAAGAAT   | CONTIONET  | ATCTTTCAAC   | CACAAATAAA                                     | TTCAAPTICA<br>AAGTPAACGT                             | GAAAATTTCA<br>CTTFFAAAGT   | NTICATTITITE TENTIFICATION AND A TENTIFICATION | CATTCAGTAG                         | TATAGCCCCA  | CCACCACATA<br>GCTGGTGTA1                                |
| 32401 | CCTTATACAG<br>CGAATATGTC                             | ATCACCOTAC   | CTTAATCAAA   | CTCACACAAC                                     | CCTACTATTC   | AACCTVATCAC<br>TTXXACGCTG  | CTCCCTCCCA   | ACACACACAGAG<br>TGTGTGTCTC         | TACACAGTCC<br>ATGTGTCAGG                          | AAAGAGGO  |
| 32501 | GCTGGCCTTA<br>CGACCGGAAT                             | ANAGCATCA  | TATCATEGGT<br>ATAGTACCCA                             | AACAGACATA<br>TTGTCTGTAT                       | TTC:TTMGGTG<br>AN:ANFCCAC                            | TTATATTCCA<br>ANTATAGGT  | CACCOTTTCC   | TGTCGAGGCA<br>ACAGCTCGGT           | AACGCTCATC  | AGTOATATI .   |
| 32601 | ATAMCTCCC<br>PATTIGAGGG                              | COOCCOTCORO  | ACTTANGTTC<br>TOAATTCAAG                             | ATGTCCCTOT<br>TACAGCGACA                       | CCAN KTROCTO<br>CCTCCACCAC                           |  | ARCCACAGGC TOCTGTCCAA<br>TUGGTGTCCG ACGACAGGTT<br>PSI  | CTTCCGOTTO<br>GAACGCCAAC           | CTTAACGGGC<br>GAATTGCCCG                          | COCCUTICOTY:  |
| 32701 | AAOTECACOC<br>TITCAGOTOCO                            | CTACATGGGG<br>GATGTACCCC                             | GTAGAGICAT<br>CATCTCAGTA                             | AATCGTGCAT<br>TTMGCACGTA                       | CAGGATAGAG<br>GTCCTATCCC                             |  | CONTRIGINATE GCAGCAGCGC  | GCGAATAAAC<br>CIXCTTATTTG          | TOCTOCCOCC<br>ACCIACGOCOG                         | OCCOCTOCOT<br>COOCGAOGCA                                |
| 32801 | CCTGCAGGAA   | PACAACATOO<br>ATOTTOTACC                             | CAGTGGTCTC<br>GTCACCAGAG<br>Pntl                     | CTCAGCGATG                                     | APPEGCACES<br>TAAGCGTRAD                             | CCCCGCACCAT<br>GCGCCGCCGTA   | AAGCCGCTT<br>TTCCCCGCAA  | OTCCTCCOOD<br>CAGGAGGCCC           | CACAGCAGCG<br>GTGTCGTCGC                          | CACCCITOAT C  |
| 32901 | TCACTTALAT<br>AGTGAATTTA                             | CAGCACAGTA   | Achecaecae<br>TOACOTEGTO                             | ACACCACAA                                      | TATTGTTCAA<br>ATAACAAGTT                             | AATCCCACAG<br>TTAGGGTGTC   | TGCAAGGCGC<br>ACGTTCCGCG   | TOTATCCAAA<br>ACATAGGTTT           | GCTCATGGCG<br>CGAGTACCGC                          | OCCACACAG<br>CCCACATGAG                                 |
| 33001 | AACCCACOFG<br>TTGGGTGCAC                             | GCCATCATAC<br>CGGTAGTATG<br>Kprii                    | CACAAGCGCA   | GCTAGATTAA<br>CCATCTAATT                       | GTOGGCTOGG<br>CACCGCTOGG                             | CTCATABACA<br>GAGTATTTGT   | CCCTUGACAT   | AAACATTACC<br>TTTOTAATOB           | TCTTTTGGCA<br>AGNAAACCGT                          | TGTTRTTAATT<br>ACAACATTAA<br>Prel                       |
| 33101 | CACCACCTCC<br>GTGGTGGAGG                             | CCCATCCATA   |  | TAAACCTCTG ATTAAACATG<br>ATTTGGAGAC TAATTTGTAC | GCGCCATCCA   | CCACCATCCT   | GCGCCATCCA CCACCATCCT NANCCAGCTG GCCAAAACCT GCCCGCCGCCGCGCGCGCGTAAGCT GGTGGTAGGA TTTGTTCATCAC CGGTTTTGGA CGGCGGCCG EGGGTAAGCT GGGCGGCCGCG EGGTTV   | GCCAAAACCT<br>COCTTTTGGA<br>EngflV | בממכנינים   | TATACACTXX:<br>ATATXTTOAC:                              |
| 33201 | AGGGAACCGG<br>TCCCTTGGCC                             | CICACCTIOT   | ATGACAGITG<br>TACTUTCACC                             | AGAGCCCAGG<br>TCTCGGGTCC                       | ACTOTTANCO<br>TOACCATTOS                             | ATCGATCATC<br>TACCTAGTAG   | ANGCTCGTCA<br>TACGAGCAGT   | TGATATCAAF                         | OTTGGCACAA<br>CAACCGTGTT                          | CACAGOCACA<br>GTGTV:CGTGT<br>PSI                        |
| 13301 | COTOCATACA   | CTTCCTCAGG<br>GAAGGAGTCC                             | ATTACAMCT<br>TANTOTTCGA                              | CCTCCCGCGT                                     | TACAACCATA<br>ATCTTAGTAT                             | TCCCAGGGAA<br>AGGGTCCCTT   | CANCCCATIC   | CTUAATCAGE                         | <b>GTAAATCCCA</b><br>CATTTAGGGT                   | CACTECAGGE  |
| 33501 | ANGACCTCOC<br>TTCTOGAGCO<br>GCAGGTAGAC<br>CCTCCATCTO | ACGTAACTCA<br>TOCATTGAGT<br>GATCCCTACT<br>CTAGGGATGA | CCTTCTCCAT<br>CCAACATCTA<br>CTACCTATTG<br>CATTCCTCAC | ACASTETICAC<br>CRECTOMINGA<br>GCGGCTETIST      | TTACATTCOS<br>ANTOTAAGCC<br>ACCCACATCG<br>TOGCTCTAGC | CHARTECTURE CONTROL TO THE THEORY OF THE THE THEORY OF THE THE THE THEORY OF THE THEORY OF THE THEORY OF THE THEORY OF THE THE THE THEORY OF THE THEORY OF THE THEORY OF THE THEORY OF THE THE THE THE THE THE THE THE THE THE | ATGATECTEC TACTAGRADO AGRETICATEC TEACAGRACO   | AOTATACCATO CAANTACATO GTTTACCTTO  | CCCCCCANG<br>OCCCCANG<br>OCCCCACOTA<br>CCCCCTCCAT | TCTCTCANA<br>ACAGAGETET F<br>GTCATATETTP<br>CAGTATAANA! |
|       |  |  |  |  | ٠  |  |  | •                                  |   |   |

# PHRKAd5gag MER682

|             |  | B.c.v. B.c.v. Hallerie   | A CALCALOR OF A | ALACATICATIVE     | מערונינינוני | Tracece: PTA | GATCCCTCTC  | TCTAGTAGTT                             | <b>GTACTATATC</b> | CACTCTCTTA    |  |
|-------------|--|--|---|-------------------|--------------|--------------|-------------|--|-------------------|---------------|--|
| 10066       | CTGROCKE                                 | MC/MADIOCO<br>Marintalogia   |   |                   |              |              | CTANCHARAC  | ACATCATCAA                             | CATCATATAG        | CTCAGAGAGT    |  |
|             | פערווייייייייייייייייייייייייייייייייייי | 200  |   |                   |              | صندسد        | TYSATAACATC | CACCACCGCA                             | GANTAAGCCA        | CACCCAGCC     |  |
| 70/65       | MAGCATCCAS                               | GEORGE IS  |   |                   |              | COCKCACCCC   | ACTATTOTAG  | progressort                            | CTTATICGGF        | GTCC(:1CO.H.  |  |
|             | 200010711                                | 2000000  | PORT PORT P   |                   |              | CAACAACCAT   |             | TTATTCCAAA                             | AGATTATCCA        | ANACCTI YANA  |  |
| 10966       | ACCTACACAT                               | TCSTRC10CG   | TO ACTUCATENCE  |                   |              | CTICTION     | CINNANNAA   | AATAAGGTIT                             | TCTAATAGGT        | TITICGAGITT   |  |
|             |  |  |   |                   |              |              |             |  |                   |               |  |
| יופטרי      | Appropriate to the second of             | Personal Personal Property Communication Com |   | and the second    | TREETCANCT   | CTACACCCAA   | AGAACAGATA  | ATCOCATTO                              | TAAGATGTTO        | CACANTOGCT    |  |
| 77301       | TACTACTAGA                               | TAATTCACTT   |   | ACCCARCCC         |              | GATCTCGGTT   | TCTTGTCTAT  | TACCGTAAAC                             | ATTCTACAAC        | groth/ccc:/   |  |
| 10076       | THE HELICAN                              | AAACCONCCT   |   | TOGACGTAAA        | CECTAMACE    | TTCAGOGTGA   | ATCTCCTCTA  | TAMCATTCC                              | ACCACCTACA        | ACCARCECTA    |  |
| 7007        | AGGITTECO                                | TTTGCCGGGA   | _   | ACCTOCATTT        | CCGATITICES  | AAGTCCCACT   | TAGACCACAT  | ATTTOTAGG                              | TCOTCOTAGG        | TOSTACOOST    |  |
| 36101       | AATAATTCTC                               | ATCTCCCCAC   | _   | TATCTCTAAG        | CAAATCCCGA   | ATATTAAGTC   | COCKCATTICT | AAAAATCTGC                             | TCCAGAGCGC        | CCTCCACCTT    |  |
|             | TTATTANGAG                               | TAGAGCGGTG   | GAAGAGITTAT   | ATAGAGATTC        | GTTTAGGGCT   | TATAATTCAG   | OCCUBITANCA | THITTAGACO                             | Ageneracoco       | GCAGGTCGAA    |  |
| 10291       | CARCCICANO                               | CACCOANTEA   | TGATTGCAMA  | AATTCAGGTT        | CCTCACAGAC   | CHCTATARIA   | TTCAMAGEG   | GAACATTAAC                             | AAAAATACCO        | CGATCCCOTA    |  |
|             | OTCOGRAPTIC                              | OPCCCTTAGE   | -   | TTAGTCCAA         | GCAGTOTCTO   | GACATATTCT   | MOTITICAL   | CHICKARITO                             | -                 | GCTRACECAT    |  |
| 34301       | CONCECUTOD                               | CAGGGCCAGC   | TGAACATAAT  | COTTICAGGTC       | TECACGRACE   | AGCGCGCGCCA  | CTICCCCCCC  | ACCAACCATG                             |                   | CCACACTOAT    |  |
|             | CCADGGAAGC                               | precedence   | ACTIGITATION  | CCACGTCCAG        | ACC/TOCC/7GO | TCGCCCCGGT   | GAAGGGGCGG  | TCCTTGGTAC                             | TOTTLETIC         | Gererania .   |  |
|             |  |  |   |                   | Hindliff     |              |             |  |                   |               |  |
| 14401       | TATRACACOC                               | ATACTCOOM  | ATACTOGGAG CTATGCTANG   | CACCCTAGCC        | CCCATCTAAG   | CTRGFTGCAT   | GGGCGGCGAT  | ATANANTOCA                             |                   | CAMMANTC      |  |
| 1           | ATACHOTOCO                               | TATORGCCTC   | GATACGATTG  | <b>OTCGCATCGG</b> | COCTACATTC   | GAACAACGEA   | CCCGCCGCTA  | TATTITACOT                             |                   | יטאדניידרנט   |  |
| 14501       | CHETAAACICCT                             | COCCCAAAAA   | NGANGCACA   | TCOTAGTCAT        | CCTCATGCAG   | NTANARGENG   | GTAAGCTCCG  | GNACCACCAC                             |                   | ACCAPITITION  |  |
| ;<br>;<br>; | CCOTTICOGA                               | OCOCOTTITI   |   | ACCATCAGTA        | CCACTACGTC   | TATTFCCGTC   | CATTCGAGGC  | CFTCCTCCTC                             |                   | Total Actions |  |
| 14601       | TCTCAAACAT                               | OPETOCOOOF   | TICTOCATAA  | ACACAMATA         | ANATAACANA   | MAACATTTA    | MCATTAGAA   | OCCIOICITA                             | -                 | AACAACCC I    |  |
|             | AGAGITTOTA                               | CAGACOCCCA   | AAGACGTATT  | TOTOTITITAT       | THATIGITY    | TITICINAL    | TTGTVATCT   | COCHCACACA                             | מדוטובביויו       | TICH COM      |  |
| 34701       | ATANGCATAA                               | GACCONCTAC   | CCCCATCCCO  | <b>GCGTGACCOT</b> | MAMANACTO    | GICACCOTCIA  | TTANAMAGCA  | CCACCCACAG                             | CHECKEGIE         | TACAGGCTC     |  |
| ·           | TATTCOTATE                               | CTOCCTOATO   | CCCCTACGCC  | CCCACTCCCA        | TEFFTTEAC    | CAGTGGCACT   | AATTTTCGT   | COLORC IOSE                            | PERCETACE         | CT.MCT.CT.NG  |  |
| 34801       | TCATANTOTA                               | AGACTCOGTA   | -   | -                 | ATCOSTCAGE   | GCTANNAGC    | GACCGAAATA  | CCCCCCCC                               | TATOTATOOG        | CONCERNATE    |  |
|             | AGTATIACAT                               | - TCTGAGCCAT   | TIGICIAOTC  | _                 | TACCCAGTCA   | CANTITICO    | _           | ************************************** | CONTRACTOR        | AATAGCACCC    |  |
| 34901       | AGACAACATT                               | ACAGCCCCCA   | -   | _                 | ATAGGAGAGA   | ANACACATA    | MACACCITION | TETTE STATE OF                         |                   | TTATCORDO     |  |
|             | <b>ACTOTIOTA</b>                         | 1 torcooods  | * ATCCTCCATA  | -                 | TATECTETE    | TTTTCTCTAT   | וופופירו    |  |                   | CACTORACA:    |  |
| 35001       | <b>PCCOGCTCCA</b>                        | GANCANCATA   | -   |                   | CCATAACAGT   | CHOCCETACC   | TCATATETE   | TTTTGGATA                              |                   | GRANCHOR      |  |
|             | Accordant                                | CFTOTTOTAT   | _   |                   |              | 76.000       | CALTABAAA   | TERRETTARCO                            | GITAMOTCC         | ACANAMARIA    |  |
| 35101       | COCACCAGET                               | CMICAGICA  | CAGTGTAMA   | MOONCCANG         | ACGRETCICET  | CATATATATE   | CTCATTTTT   | ACTOCATTOC                             | -                 | TOTTTTOT      |  |
|             | CCOTOCACOA                               | ומנושמום   |   |                   | CLARABARC    | CACAACTTCC   | TCAAATCGTC  | ACTICCOTT                              | TCCCACGTTA        | CONCACTICC    |  |
| 35201       | GCCAGAAAAC                               | COCACOCOAN<br>COCACOCOAN   | CCTACOCCCA<br>CCATCCCCCT  |                   | GCTTTTT00    | CICITICAAGG  | -           | TGAAGCCAAA                             | AGGGTGCANT        | GCAGTGAAGG    |  |
|             |  |  |   |                   |              |              |             |  |                   |               |  |

Figure ISV

# PHRKArlSgag MER682

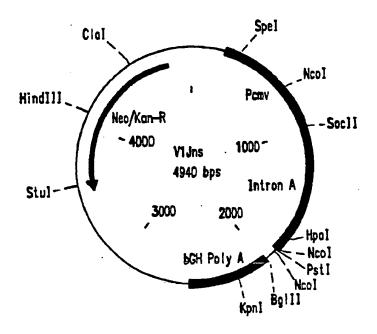
| 35301  | CATTITANDA        | CANTITARDA AARCTACAAT TECEAACACA TACAAGITAC TECAGECAAA AARCTACATE AEEGGEEEGG ITEECAAGEE EGGGECAEG TEACAAAFTE | TOCCANCACA         | TACAAGTTAC            | Trenceta   | ACCTACCTIC   | ACCOGCCCC                               | TICCCANGCC        | CCCCCCACG          | TCACAMCTC     |
|--------|-------------------|--|--------------------|-----------------------|--|--------------|---|-------------------|--------------------|---------------|
|        | <b>OTAMATHCP</b>  | P TITIGATGITA  | AGGGTTCTGT         | ATCTTCAATG            | MACCHAINT  | TTGGATOCAG   | TTGGATGCAG TGAGGGGGGC                   | AGGGTGCGG         | <b>OCCOCOOTI</b>   | ACTICITITIONS |
| •      |                   |  |                    |                       |  | Pac          | ed<br>mmmm<br>Frefil                    |                   | ٠                  |               |
| 15.404 |                   |  |                    |                       |  |              | *************************************** |                   |                    |               |
| 1010   | GTTAGGGGAGT       | C AATROTESTAA  | CCGAAGTTAG         | CAMMATANCE            | ATATAATAA  | TACTACANTE   | AATTCTAAG                               | CCTAGACGCT        | GCGCAGGCTG         | CTACCOCATA    |
|        |                   |  |                    |                       |  |              |   |                   |                    |               |
| 19666  | CCCATTATOA        | THE PREPERCY   |                    | TAGGGGATAGG           | CCCCCTTGCA   | CYCCATGCTO   | ACCAGGGAGG                              | ATTACTACGA        | CCATCAGGGA         | CAGCTTCAMD    |
| 15601  | CTTREETABLE       | •  | _                  | _                     | -  | CATACATITUT  | GULLERA                                 | CONTRACTOR        | ANDARTICAL         | CUMP. A BUTTE |
|        | COOLCOTTIT        | _  | _                  | _                     | -  | _            | CCCCCCACT                               | OCTOGRAGIO        | Trritacto          | CONCITICACT   |
| 35701  | CACCTOCCCA        |  | _                  |                       |  | CAACCTCCCT   | CONCOCICT                               | CCTUTTCCGA        | CICTOCCOCT         | TACCCOATAC    |
|        | CTCCACCOCT        | •  | _                  | •                     | _  |              | CCACCCCAGA                              | COACAAGGCT        | CONACCOCCA         | ATORCCTATA:   |
| 35801  | CTUTCCOCCT        | r rececenc   | GOGAAGCHTO         | acactracte            | ATARCTCACG   | CTOTACACTAT  | CYCANTICOS                              | TOTAGGTCOF        | TECCTECANO         | CTOCOCCTOD.   |
|        | GACAGGCGGA        | ANDROGGANG   | CCCTTCGCAC         | COCCANACING           | TATCANGICC   | GACATCCATA   | GATTCANGC                               | ACATCCAGCA        | AGCGAGGTTC         | GACCCGACA!    |
| 35901  | TOCACOAACC        | : CCCCGTTCAG   | CCCGACCGCT         | <b>GCGCCTTATC</b>     | CCCTANCTAT   | CONCINGAGE   | CCANCECOST                              | AAGACACGAC        | Trateoccae         | тоослослог    |
|        | ACCTOCITION       | -  | OGGETGGCCA         | COCOGNATAG            | <b>OCCAPTGATA</b>  | CCAGAACTCA   | GCTTOGGCCA                              | THETETOCTO        | <b>NATABCCOOTO</b> | ACCGICCOTC    |
| 36001  | CACTOOTAAC        | : AGGATTAGCA   | GAGCGAGGTA         | TOTAGCCOOT            | CCTACACACT   | TITTGAAGTO   | STOCCCTMC                               | TACOGCTACA        | CTAGANGGAE         | AGTATTIÓGE    |
|        | OTOACCATTO        | J TECTAATEGE   | CICGCICCAT         | ACATCCCCCA            | CGATUTCTCA   | AGNACTITCAC  | CACCOGATTO                              | ATOCCOATOT        | DATETTECTO         | TCATAAACCA    |
| 36101  | ATCTOCGCTC        | : TOCTOMOCC  | Agtraccric         | CONTANACAG            | THESTAGETE   | TrGATCCGGC   | ANACAAACCA                              | cccctootho        | COORGOTTIT         | TTTGTTTGC.    |
|        | TAGACGCGAG        | ACCACTTCGG   | <b>f</b> CANTGGANG | commence              | NACCATCCAG   | AACTAGGCCG   | THEFTOOT                                | <b>GOCGACCATC</b> | OCCYCCANA          | AMCARACTT     |
| 36201  | AGCAGCAGAT        | TACGCGCAGA   | AAANANGGAT         | CTCAAGAAGA            | TCCTTTGATC   | TITITICTACGO | GCTCTGACGC                              | TCAGTOGNAC        | GANAACTCAC         | GTTANOCKA     |
|        | <b>ACOTOOTETA</b> | ATOCOCONCT   | TITITICCTA         | GAGITICITICE          | ACCINACTAG   | ANAGATOCC    | CCAGACTOCO                              | AGICACCITIO       | CTTHOAGTO          | CAAINCCCTA    |
| 36301  | TITOGREATO        | AGATTATCAA   | AAAGGATCTT         | CACCTAGATO            | CITTIAMIC  | ANTCHANGE    | ATATATOAGT                              | AAACTTICGTIC      | TOACAGTTAC         | CNATICETER    |
|        | AAACCAGTAC        | TCTANTAGET   | THICCTAGA          | GTCCATCTAG            | GAAAATTTAG   | TTAGATTTCA   | TATATACTCA                              | THTCANCCAG        | ACTOTCAATO         | CTTACCAMATO   |
| 36401  | TCAGTGAGGC        | ACCTATCTCA   | OCCUNICACIO        | TATTICOTIC            | ATCCATAGET   | OCCIONCICC   | CCOTCOTOTA                              | GATAACTACG        | ATACOGGAGG         | OCTTACCATY:   |
|        | AGTCACTCCO        | TOGATAGAGE   | CGCTAGACAG         | ATAAAGCAAG            | TACCTATCAA   | CCAGACTICAGO | <b>GGCAGCACAT</b>                       | CTATTGATGC        | TATOCCCTCC         | CCMATCISTAG   |
| 36501  | TOOCCCCAOT        | <b>GCTOCAATOA</b>  | TACCOCCAGA         | CCCACGCTCA            | CCGCCTCCAG   | ATTTATCAGE   | ANTANACCAG                              | CCACCCCCAA        | GCCCCCACCG         | CACAMOTOGT    |
|        | ACCOGGGTCA        | COACGITACT   | ATGGCGCTCT         | <b>GOGTOCCANT</b>     | CCCCACATC  | TANATAGECT   | TTATTTGGTC                              | gorcoocc11        | CCCOGCICGC         | GICTICACCA    |
| 36601  | CCTGCACTT         | . PATCCOCCTC   | CATCCAOTCT         | ATTAATTICIT           | <b>accessonac</b>  | TAGASTAAST   | AGITCOCCAG                              | TTRATACTIT        | GCGCAACGTT         | GITTCCCATTG   |
|        | OGACOTICAA        | _  | _                  | TAATTAACAA            | COCCCTTCG  | ATCTCATTCA   | TEAMGEOUTE                              | ANTTATCAMA        | COCOTTOCAA         | CANCOGTAAC    |
| 36701  | CTACAGOCAT        | COTOGRACA  | COCTOGREGE         | TROOTATOOC            | TTCATTCAGC   | Teccorrece   | AACCATCAAG                              | GCGAOTTACA        | TCATCCCCCA         | TOTTOTOTA     |
|        | GATOTCCOTA        | _  | OCCUOCACIA         | AACCATACCO            | AAGTAAGTCO   | ACCCANGCG    | TRECTACTIC                              | COCTCAATOT        | ACTAGGGGGT A       | ACAACACGTT    |
|        |                   |  | Ž                  |                       |  |              |   |                   |                    |               |
| 36801  | AAAAGCOGIT        | AGCTCCTTCO   | GTCCTCCOAT         | Greenecoar correspond | ACTAACTING   | CCCCAGTGTT   | NTCACTCATE                              | <b>CITATOCCAG</b> |                    | TICACTIACT    |
| •      | TTTTCGCCAA        | ٠.   |                    | <b>GCAACAGTICT</b>    | TCATTCACC  | COCCITCACAA  | TACHGAGTAC                              | CANTACCOTC        | OTORCGTATT !       | AACAGAATTA    |
| 16901  | GTCATGCCAT        | CCUTANGATO   | CTTTTCTGTG         | ACTRICATER            | ACTUACCAA  | GICATICIOA   | GAATAGTGTA                              | TOUNGCOACE        | GARTIGGERCT        | Tracconner    |
| 1      |                   |  |                    |                       | The Production of the Producti | CACTBACACT   | CTTATTACAT                              | ACCICCACTOG       | CTCNACGAGA         | ACCICICCO     |

Figure 15W

## PMRKAd5gag MRR682

|   | ID NO: 27)   | TAT (SEQ    | GATGACATTA ACCTATAAAA ATAGGCGTAT GACGAGGCCC TITICGTCTTC AAGAATTGGA TITICGAATTCT TAAT (SEQ ID NO: 27) GTACTGTAAT TGGATTTTT TATICGCATA GTGCTCCTGG AAAGGAGAAG TICTTAACT AGGCTAAAA ATTA (SEQ ID NO: 28) | Barrell<br>ANGANTITGA TU | TTICGICTIC    | CACGAGGCCC   | ATACGCGTAT<br>TATCCGCATA | ACCTATAAAA<br>fogatatiti | CATOACATTA<br>GTACTOTAAT | 37401 |
|---|--------------|-------------|---|--------------------------|---------------|--------------|--------------------------|--------------------------|--------------------------|-------|
| COCCENIOTA TARACTERCA TRANSCIENT TATFICITIN ICCCCANAX GCGTGTANG GGGCTTITCA COGTGGACTO CAGATICITI GGTAATAATA | CAGATTETT    | COCTODACTO  | GOCCHTTICA C  | GCCTCTANG                | Teerchara     | TATTICITIA   | TAMTETITE                | TARACTTACA               | COCCTATOTA               |       |
| CCATTATTA   | GTCTANGAM    | CCACCTOAC   | CCCGANAGT O   | CLEVENTATIVE             | ACCIPATING    | ATAMCANT     | ATTTAGAMAA               | ATTICAATOT               | GCCGATACAT               | 37301 |
| ACAGAGTAL"  | AGTOCCANTA   | TECOTAAAT   | CTTATAATAA C  | GAAGTIAAAAA              | ATTCACT'ATT'A | TTTACAACTT   | CCGCTRITICC              | <b>TCCCTTATTC</b>        | COOCUMENT                |       |
| TCTCTCATCA  | TCACKAGITTAT | MAGCATTTA   | CAATATTATT G  | rmccmm                   | TACTUATACT    | ANATICITICAN | COCCACCOG                | ACCENATION               | <b>GCCGCANAA</b>         | 37201 |
| THECHTITA   | gritteree    | CACCCACTC   | CTCCTCCCAA A  | CANATTEMA                | ACAACTCCTA    | Tracerrence  | COTTCACCACG              | ACCTACATTO               | CTCTAGGTCA               |       |
| MUCCUMANT   | CAANAACNOO   | CTCCCTCAG   | CACCASICIST 1   | TTTAC:TTT.               | TKTEN:NTT     | ACTEANETES   | CCACTCGTGC               | TCGATGITANC              | DADATECAGE               | 37101 |
| ATCCCCACA   | NOTICCTAGA   | OCTITITIONS | באעמאסטכם ם   | TAACCITITIG              | Tr:Accaletan  | CTICAAATT    | OCTOTATEGE               | ATTATOCCCC               | <b>OPTOTOCCCT</b>        |       |
| בשרנו זיר וזיו ז  | יייינאטעייי  | COMMACIC    | CHICAMORE C   | ATTRIBUTE                | ACTIVITY ATT  |              | CCACATARCA               | TANTACCOCO               | CACACACAGA               | 37001 |

Figure 15X



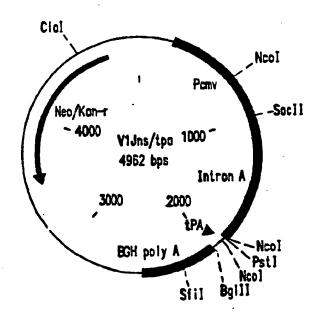


FIGURE 16

GCAGTGGCCCCTGACTGAGGAGAAGATCAAGGCCCTGCTGGAAATCTGCACTGAGATGGAGAAGGAGGGCAAAATCTCCA sGinTrpProLeuThrGiuGiu\_yslieLysAioLeuVoiGiuIieCysThrGiuMelGiuLysGiuGiyLyslieSerL 30 40 50

ACATTGCCCCGGAGAACCCCTACAACACCCCTGTGTTTGCCATCAAGAAGAAGAAGGACTCCACCAAGTGGAGGAACCTGGTG
ysleGlyProGluAsnProTyrAsnThrProVolPheAlolleLysLysLysAspSerThrLysTrpArgLysLeuVol
60 70

GACTTCAGGGAGCTGAACAAGAGGACCCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCCACCCCGCTGGCCTGAAGAA AspPheArgGTuLeuAsnLysArgThrGTnAspPheTrpGTuVoTGTnLeuGTyTleProHisProAloGTyLeuLysLy 80 90 100

GAAGAAGTCTGTGACTGTGGCTGGCGGGATGCCTACTTCTCTGTGCCCCTGGATGAGGACTTCAGGAAGTACACTG slyslysSerVolThrVolLeu<u>Alo</u>VolGlyAspAloTyrPheSerVolProLeuAspG1uAspPheArgLysTyrThrA 110 120 130

CCTTCACCATCCCCTCCATCAACAATGASACCCCTGGCATCAGGTACCAGTACAATGTGCTGCCCCAGGGCTGGAAGGGC loPheTnrlleProSerlleAsnAsnGluThrProGlylleArgTyrGlnTyrAsnVolLeuProGlnGlyTrpLysGly 140 150

TCCCCTGCCATCTTCCAGTCCTCCATGACCAAGATCCTGGAGCCCTTCAGGAAGCAGAACCCTGACATTGTGATCTACCA SerProAlollePheGInSerSerMetThrLyslleLeuGluProPheArgLysGInAsnProAsplleVollleTyrG1 160 170 180

TGCTGAGGTGGGGCCTGACCACCCCTGACAAGAAGCACCAGAAGGAGCCCCCCTTCCTGTGGATGGGCTATGAGCTGCAC euleuArgTrpGTyLeuThrThrProAsplysLysHisGInLysGTuProProPheLeuTrpMetGTyTyrGTuLeuHis 220 230

DCCGACAGTGGACTGTGCACCCATTGTGCTGCCTGAGAAGGACTCCTGGACTGTGAATGACATCCAGAAGCTGGTGGG ProAsplystrpThrVoIGinProIieVoileuProGiulysAspSerTrpThrVoiAsnAspIieCinlysLeuVoiGI 240 250 260

CAAGCTGAACTGGGCCTCCCAAATCTACCCTGGCATCAAGGTGAGGCAGCTGTGCAAGCTGCTGAGGGGCACCAAGGCCCC yLysLeuAsnTrpAloSerGinlieTyrProGiylieLysVoiArgGinLeuCysLysLeuLeuArgGiyThrLysAloL 270 280 290

#### FIGURE 17A

TGACTGAGGTGATCCCCCTGACTGAGGAGGCTGAGCTGGAGCTGGGCTGAGAACAGGGAGATCCTGAAGGAGCCTGTGCAT EUThrG1uVo111eProLeuThrG1uG1uA1oG1uLeuG1uLeuG1uLeuA1oG1uAsnArgG1u11eLeuLysG1uProVo1His 300 310

GCCGTGTACTATGACCCCTCCAAGGACCTGATTGCTGAGATCCAGAAGCAGGGCCAGGGGCCAGTGGACCTACCAAATCTA GiyVoiTyrTyrAspProSerLysAspLeuiteAtoGtuiteGinLysGtnGtyGtnGtyGtnTrpThrTyrGtn1teTy 320 340

CCADGAGCCCTTCAAGAACCTGAAGACTGCCAAGTATCCCAGGATGAGGGGGGCCCCACACCAATGATGTGAAGCAGCTGA rGinGluProPheLysAsnLeuLysThrGlyLysTyrAloArgMetArgGlyAloHisThrAsnAspVolLysGInLeuT 350 350 370

CICAGGCTGTGCAGAAGATCACCACTGAGTCCATTGTGATCTGGGGCAAGACCCCCAAGTTCAAGCTGCCCATCCAGAAG hrGluAloVolGinLyslieThrThrGluSerlieVollieTrpGlyLysThrProLysPheLysLeuProlleGinLys 380 390

GGTGAAGCTGTGGTACCAGCTGGAGAAGGAGCCCCATTGTGGGGGGCTGAGACCTTCTATGTGGCTGGGGCTGCCAACAGGG uVollysleu1rpTyrGinleuGiuLysGiuProlleVolGlyA1oGiuThrPheTyrVolA1oGiyA1oAioAsnArgG 430 440 450

AAGACTGCCCTCCAGGCCATCTACCTGGCCCTCCAGGACTCTGGCCTGGAGGTGAACATTGTGACTGCCTCCCAGTATGC
LysThrAloLeuGinAlolleTyrLeuAloLeuGinAspSerGlyLeuGluVolAsnIleVolThrAloSerGInTyrAl
480
490
500

CCTCGGCATCATCCACGCCCAGCCTGATCAGTCTGACTCTGACCTGGTGAACCAGATCATTGAGCAGCTGATCAAGAAGG oLeuGiylielleGinAioGinProAspGinSerGiuSerGiuLeuVolAsnGinIielleGiuGinLeuIielysLysG 510 520 530

AGAAGGTGTACCTGCCCTGCCCACAAGGCCATTGCGGGCAATGAGCAGGTGGACAAGCTGGTGTCTGCTGGC
IULysVolTyrLeuAloTrpVolProAloHisLysGiyileGiyGiyAsnGiuGinVolAspLysLeuVolSerAloGiy
540
550

ATCAGGAAGGTGCTGTTCCTGGATGGCATTGACAACCCCCAGGATGAGCATGAGAACTACCACTCCAACTGGAGGGCTAT
11eArglysVolleuPheleuAspClyI1eAsplysA1oG1nAspG1uHisG1uLysTyrHisSerAsnTrpArgA1oMe
560 570 580

#### FIGURE 17B

GCCCTCTGACTTCAACCTGCCCCCTGTGGTGGCTAAGGAGATTGTGGCCTCCTGTGACAAGTGCCAGCTGAAGGGCGAGG tAloSerAspPheAsnLeuProProVolVolAloLysGlulleVolAloSerCysAspLysCysGlnLeuLysGlyGluA 590 600 610

GCTGTGCATGTGGCCTCCGGCTACATTGAGGCTGAGGTGATCCCTGCTGAGACAGGCCAGGAGACTGCCTACTTCCTGCT AlovolHisVolAloSerGlyTyrlleGluAloGluVollleProAloGluThrGlyGlnGluThrAloTyrPheLeuLe 640 650 660

GAAGCTGGCTGGCAGGTGGCCTGTGAAGACCATCCACACTGCCAATGGCTCCAACTTCACTGGGCCCACAGTGAGGGCTG
uLysLeuAldGlyArgTrpProVolLysThrIleHisThrAloAsnGlySerAsnPheThrGlyAloThrVolArgAloA
680
690

CCTGCTGGTGGCCTGGCATCAASCAGGAGTTTGGCATCCCCTACAACCCCCAGTCCCACGGGTGGTGGCCTCCATGAAC
IOCysTrpTrpAioGlyIleLysGInGluPheGlyIleProTyrAsnProGInSerGInGlyVolVolAlaSerMelAsn
700 710

AAGGAGCTGAAGAAGATCATTGGGCAGGTGAGGGACCAGCTGAGCACCTGAAGACAGCTGTGCAGATGGCTGTGTTCAT LysG1uLeuLysLys11e11eG1yG1nVo1ArgAspG1nA1oG1uHisLeuLysThrA1oVo1G1nMetA1oVo1Phe11 720 730 740

CCACAACTTCAAGAGGAAGGGGCGCATCGGGGGCTACTCCGCTGGGGAGAGGATIGTGGACATCATTGCCACAGACATCC
eHisAsnPhelysArglysGlyGlylleGlyGlyTyrSerAloGlyGluArglleVolAsplleIleAloThrAsplleG
750
770

AGACCAAGGAGCTCCAGAAGCAGATCACCAAGATCCAGAACTTCAGGGTGTACTACAGGGACTCCAGGAACCCCCTGTGG
InThrLysGIuLeuGInLysGInlieThrLysIieGInAsnPheArgVoITyrTyrArgAspSerArgAsnProLeuTrp
780 790

AAGGCCCTGCCAAGCTGCTGTGGAAGCGGGAGCGGGTGTGGTGATCCAGGACAACTCTGACATCAAGGTGGTGCCCAG LysGtyProAtoLysLeuleuTrpLysGtyGtuGtyAtoVotVotIteGtnAspAsnSerAsp1teLysVotVotProAr 800 810 820

AMACCCCCCCCAGATCT (SEQ ID NO: 3)
Xx Bq 11 (SEQ ID NO: 4)

FIGURE 17C

CATCACCATGCATGCAATGAAGAGAGGCTCTGCTGTGCTGCTGCTGTGTGGAGCAGTGTTTGCC

MetaspalometlysargGjyleuCysCysVolleuLeuleuCysGjyaloVolPheVaiSerP

-25

(within SEQ 10 NO: 7) (within SEQ 10 NO: 8) RoSerGiulieSerAidProlieSerProlieGiuThrVolProValLysLeuLysProGlyMelAspGly 20

#### FIGURE 18

| WT   | - ATG GGT GGC AAG TGG TCA AAA CGT AGT GTG CCT GGA TGG TCT                                | -42          |
|------|--|--------------|
| OPT  | M G G K W S K R S V P G W S  | -14          |
| ₩T   |  | -84          |
| OPT  | - ÁCC GTG ÁGG GÁG ÁGG ÁTG ÁGG AGG GCC GAG CCC GCC GCC GAC<br>T V R E R M R R A E P A A D | -28          |
| WT.  | - AGG GTG AGA CGA ACT GAG CCA GCA GCA GTA GGG GTG GGA GCA                                | -126         |
| OPT  | - ÁĞĞ ĞTĞ ÁĞG AĞG ÁCC ĞÁĞ CCC ĞCC ĞCC ĞTĞ ĞĞC ĞTĞ ĞĞC ĞCC<br>R V R R T E P A A V G V G A | -42          |
| WT   | - GTA TCT CGA GAC CTG GAA AAA CAT GGA GCA ATC ACA AGT AGC                                | -168         |
| OPT  | - GTG TCC AGG GAC CTG GAG AAG CAC GGC GCC ATC ACC TCC TCC V S R D L E K H G A I T S S    | -56          |
| WT   | - AAT ACA GCA GCT ACC AAT GCT GAT TGT GCC TGG CTA GAA GCA                                | -210         |
| OPT  | - AAC ACC GCC GCC ACC AAC GCC GAC TGC GCC TGG CTG GAG GCC<br>N T A A T N A D C A W L E A | <b>-7</b> 0. |
| WT . | - CAA GAG GAT GAG GAA GTG GGT TTT CCA GTC AGA CCT CAG GTA                                | -252         |
| OPT  | - CAG GAG GAC GAG GAG GTG GGC TTC CCC GTG AGG CCC CAG GTG Q E D E E V G F P V R P Q V    | -84          |
| WT   | - CCT TTA AGA CCA ATG ACT TAC AAG GGA GCT GTA GAT CTT AGC                                | -294         |
| OPT  | - CCC CTG AGG CCC ATG ACC TAC AAG GGC GCC GTG GAC CTG TCC PLRPMTYKGAVDLS                 | -98          |
| WT   | - CAC TIT TTA AAA GAA AAG GGG GGA CTG GAA GGG CTA ATT CAC                                | -336         |
| OPT  | - CAC TTC CTG AAG GAG AAG GGC GGC CTG GAG GGC CTG ATC CAC H F L K E K G G L E G L I H    | -112         |
| WT   | - TCA CAG AAA AGA CAA GAT ATC CTT GAT CTG TGG GTC TAC CAC                                | -378         |
| OPT  | TCC CAG AAG AGG CAG GAC ATC CTG GAC CTG TGG GTG TAC CAC S Q K R Q D I L D L W V Y H      | -<br>-126    |
| WT   | - ACA CAA GGC TAC TTC CCT GAT TGG CAG AAC TAC ACA CCA GGG                                | -420         |
| OPT  | - ACC CAG GGC TAC TTC CCC GAC TGG CAG AAC TAC ACC CCC GGC T Q G Y F P D W Q N Y T P G    | -140         |

FIGURE 19A

| WT         | - CCA GGA ATC AGA TTT CCA TTG ACC TTT GGA TGG TGC TTC AAG -462                                    |          |
|------------|---|----------|
| OPT        | - CCC GGC ATC AGG TTC CCC CTG ACC TTC GGC TGG TGC TTC AAG<br>P G I R F P L T F G W C F K -154     |          |
| WT         | - CTA GTA CCA GTT GAG CCA GAA AAG GTA GAA GAG GCC AAT GAA -504                                    |          |
| OPT        | - CTG GTG CCC GTG GAG CCC GAG AAG GTG GAG GAC GAG GCC AAC GAG<br>L V P V E P E K V E E A N E -168 |          |
| WT         | - GGA GAG AAC AAC TGC TTG TTA CAC CCT ATG AGC CAG CAT GGG -546                                    |          |
| <b>DPT</b> | - GẬC GÁG ÁÁC ÁÁC TỚC CTỔ CTỔ CÁC CÓC ÁTG TCC CÁG CÁC GỚC<br>G E N N C L L H P M S Q H G -182     |          |
| WT         | - ATA GAG GAC CCG GAG AAG GAA GTG TTA GAG TGG AGG TTT GAC -588                                    | ;        |
| OPT        | - ÁTC GÁG GÁC CĆC GÁG ÁÁG GÁG GTG CTG GÁG TGG ÁGG TTC GÁC<br>I E D P E K E V L E W R F D -196     | <b>;</b> |
| WT .       | - AGC AAG CTA GCA TTT CAT CAC GTG GCC CGA GAG CTG CAT CCG -630                                    | )        |
| OPT        | TCC AAG CTG GCC TTC CAC CAC GTG GCC AGG GAG CTG CAC CCC  S K L A F H H V A R E L H P -210         | )        |
| WT         | - GAG TAC TAC AAG GAC TGC TGA (SEQ ID NO:30) -65  | ì        |
| OPT        | - GAG TAC TAC AAG GAC TGC TAA (contained within SEQID NO: 9) E Y Y K D C (SEQID NO: 10) -21       | 5        |

FIGURE 19B

VIJns/nef

CARGGSTCTTTTCTGGGGGCTCCTTGGGGTCTGCCACC ATG GGC GGC ANG TGG TCC ANG NGG TCC GTG CCC .

M G G K W S K R S V P

Srf1 Ball1 . . . . CAC CCC GAG TAC TAC AAG GAC TGC TAA AGCCCGGGAJCTGCTGCCTTCTAGTTGCCAGC (SEQ 1D NU: 38) H P E Y Y K D C \* (contained within SEQ 1D NO: 10:

V1Jns/nef(G2A.LLAA)

Psti Catrogetett transferencestect transference of the GCC GGC AAG TGG TCC AAG AGG TCC GTG CCC .

M A G K W S K R S V P

Srff BOILT

CAC CCC GAG TAC TAC AAG GAC TGC TAA AGCCCGGGCACAICTGCTGTGCCTTCTAGTTGCCAGC (SEQ 10 NO: 39)

H P E Y Y K D C \* (Contained Wilhin SEQ ID NO:14)

Vljns/tpanef & Vljns/tpanef(LLAA)

*Pst1 Catgggictiggggicaccgticttatatictagaticacc* atg gat gca atg ang aga ctc tgc tgt gtg m d a m k r g l c c v

CTG CTG CTG TGT GGA GCA GTC TTC GTT TCG CCC AGC GAG AIC ICC TCC AAG AGG TCC GTG CCC ...

. . CAC CCC GAG TAC TAC AAG GAC TGC TAA AGCCCCGGGAGATCTGCCTCTCTAGTTGCCAGC (SEQ ID NG: 40)

H P , E Y Y K D C \* (contained withon SEQ ID NO: 16)

## FIGURE 20

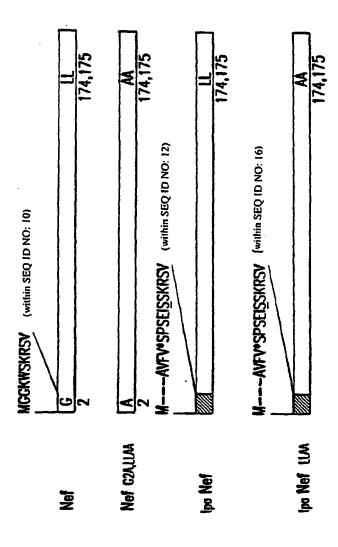


FIGURE 21

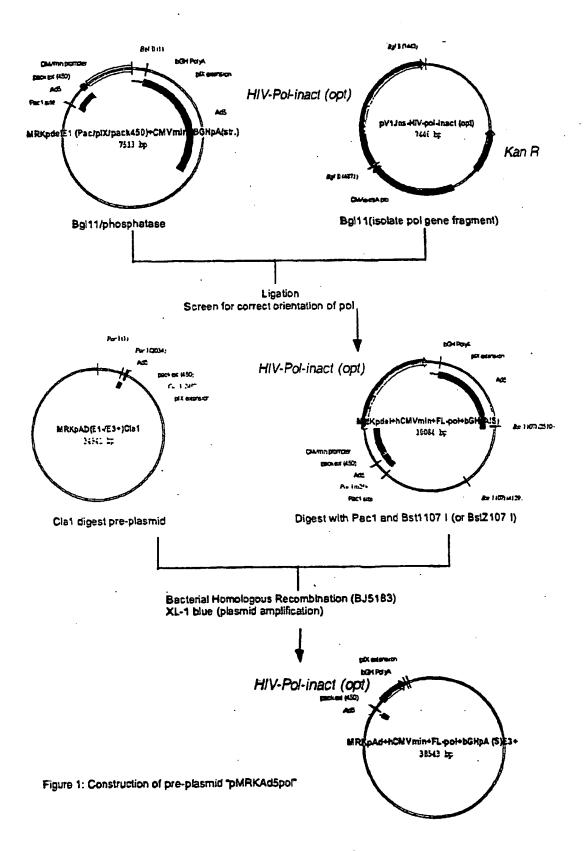


FIGURE 22

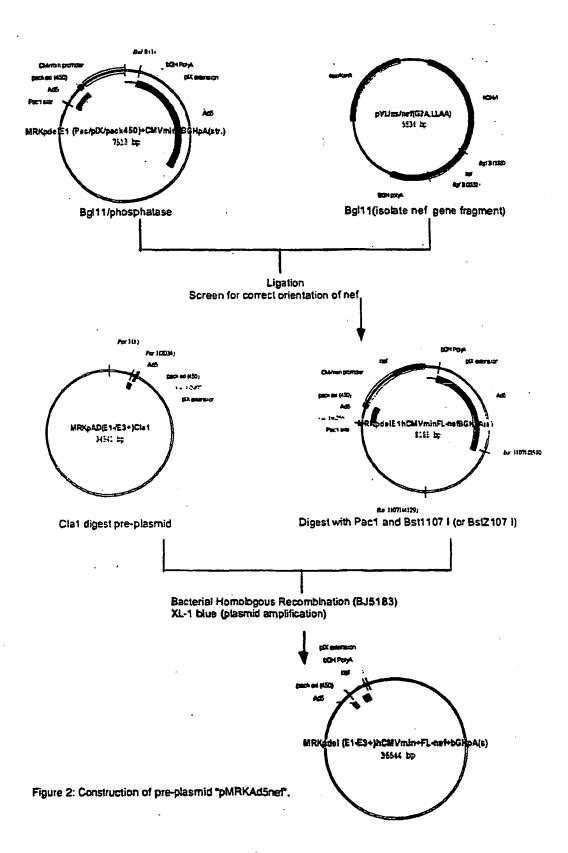


FIGURE 23

Comparison of Clade B vs. Clade C Anti-gag T Cell Responses in Clade B HIV-Infected Subjects

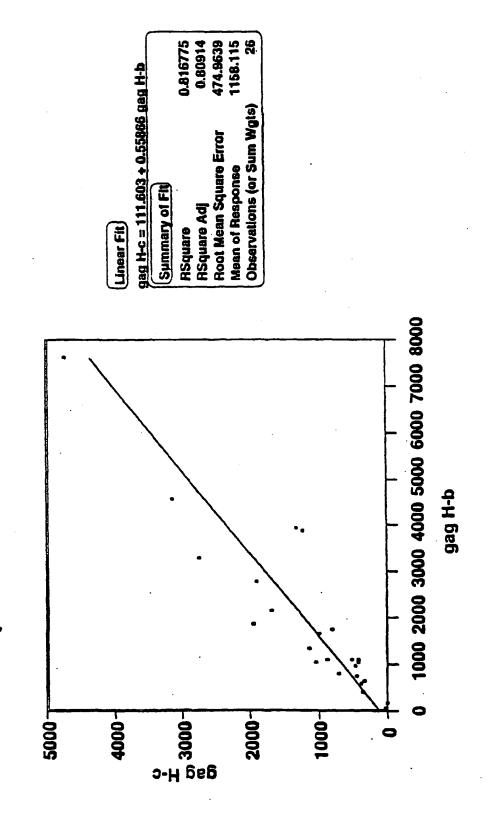
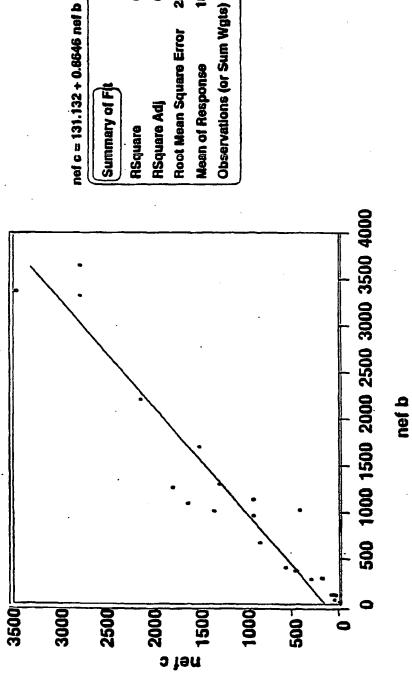


FIGURE 25

289.7718 1096.435 R

0.91685

Comparison of Clade B vs. Clade C Anti-nef T Cell Responses in Clade B HIV-Infected Subjects



## MRKAd5pol MER1062 (MRKAd5 Pre-Adenoviral Vector Containing the IA opt pol Coding Region)

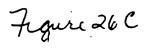
| 1   | CATCATCAAT |              |            |              |  |
|-----|------------|--------------|------------|--------------|--|
|     | GTAGTAGTTA | TTATATGGAA   | TAAAACCTAA | CTTCGGTTAT   | ACTATTACTC                               |
| 51  | GGGGTGGAGT | TTGTGACGTG   | GCGCGGGGCG | TGGGAACGGG   | GCGGGTGACG                               |
|     |            |              | CGCCCCCCC  |              |  |
| 101 |            |              |            |              | ACACATGTAA                               |
|     |            |              | CTACAACGTT |              |  |
| 151 |            |              | GACGTTTTTG |              |  |
|     |            |              | CTGCAAAAAC |              |  |
|     | GAAGTGACAA |              |            |              |  |
|     | CTTCACTGTT |              |            |              |  |
| 251 | CGTAACCGAG | TAAGATTTGG   | CCATTTTCGC | GGGAAAACTG   | AATAAGAGGA                               |
|     |            |              |            |              | TTATTCTCCT                               |
| 301 | AGTGAAATCT | GAATAATTTT   | GTGTTACTCA | TAGCGCGTAA   | TATTTGTCTA                               |
|     |            |              |            |              | ATAAACAGAT                               |
| 351 |            |              |            |              | AGGTGTTTTT                               |
|     |            |              |            |              | TCCACAAAAA                               |
| 401 |            |              |            |              | ATTATTATAG                               |
|     |            | •            |            |              | TAATAATATC                               |
| 451 |            |              |            |              | TATGTACATT                               |
|     |            |              |            |              | ATACATGTAA                               |
| 501 | TATATTGGCT | CATGTCCAAC   | ATTACCGCCA | TGTTGACATT   | GATTATTGAC                               |
|     |            |              |            |              | CTAATAACTG                               |
| 551 |            |              |            |              | AGCCCATATA                               |
|     |            |              |            |              | TCGGGTATAT                               |
| 601 | TGGAGTTCCG | CGTTACATAA   | CTTACGGTAA | ATGGCCCGCC   | TGGCTGACCG                               |
|     |            |              |            |              | ACCGACTGGC                               |
| 651 |            |              |            |              | TTCCCATAGT                               |
|     |            |              |            |              | AAGGGTATCA                               |
| 701 | AACGCCAATA | GGGACTTTCC   | ATTGACGICA | ATGGGTGGAG   | TATTTACGGT                               |
|     |            |              |            |              | ATAAATGCCA                               |
| 751 | AAACTGCCCA | CTTGGCAGTA   | CATCAAGTGT | ATCATATGCC   | AAGTACGCCC                               |
|     |            |              |            |              | TTCATGCGGG                               |
| 801 | CCTATTGACG | TCAATGACGG   | TAAATGGCCC | C GCCTGGCATT | ATGCCCAGTA                               |
|     |            |              |            |              | TACGGGTCAT                               |
| 851 | CATGACCTTA | 1 TEEGACTITY | CTACTTGGC  |              | GTATTAGTCA                               |
|     |            |              |            |              | 10 10 10 10 10 10 10 10 10 10 10 10 10 1 |

7 i jure 26A

| 901  |                          | CATGGTGATG<br>GTACCACTAC |      |  |
|------|--------------------------|--------------------------|------|--|
| 951  |                          | ACTCACGGGG<br>TGAGTGCCCC |      |  |
| 1001 |                          | TTTTGGCACC<br>AAAACCGTGG |      |  |
| 1051 |                          | CCCATTGACG<br>GGGTAACTGC |      |  |
| 1101 |                          | GCAGAGCTCG<br>CGTCTCGAGC |      |  |
| 1151 |                          | TGTTTTGACC<br>ACAAAACTGG |      |  |
| 1201 |                          | GGAACGGTGC<br>CCTTGCCACG | <br> |  |
| 1251 |                          | ATGGCCCCCA<br>TACCGGGGGT |      |  |
| 1301 |                          | CATGGATGGC<br>GTACCTACCG |      |  |
| 1351 |                          | AGGCCCTGGT<br>TCCGGGACGA |      |  |
| 1401 |                          | AAGATTGGCC<br>TTCTAACCGG | ·    |  |
| 1451 |                          | GAAGGACTCC<br>CTTCCTGAGG |      |  |
| 1501 |                          | AGAGGACCCA<br>TCTCCTGGGT | <br> |  |
| 1551 |                          | GGCCTGAAGA<br>CCGGACTTCT |      |  |
| 1601 | GGGATGCCTA<br>CCCTACGGAT | CTTCTCTGTG<br>GAAGAGACAC |      |  |
| 1651 | GCCTTCACCA<br>CGGAAGTGGT | TCCCCTCCAT<br>AGGGGAGGTA |      |  |
| 1701 | GTACAATGTG<br>CATGTTACAC | CTGCCCCAGG<br>GACGGGGTCC | <br> |  |
|      | CCTCCATGAC<br>GGAGGTACTG |                          | <br> |  |
| 1801 | GTGATCTACC<br>CACTAGATGG | AGTACATGGC<br>TCATGTACCG |      |  |



| 1901 | GGGGCCTGAC<br>CCCCGGACTG | CACCCCTGAC<br>GTGGGGACTG |                          |                          |                          |
|------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| 1951 |                          | ATGAGCTGCA<br>TACTCGACGT |                          |                          |                          |
| 2001 | ••••                     | AAGGACTCCT<br>TTCCTGAGGA |                          |                          |                          |
| 2051 |                          | CTGGGCCTCC<br>GACCCGGAGG |                          |                          |                          |
| 2101 |                          | TGCTGAGGGG<br>ACGACTCCCC |                          |                          |                          |
| 2151 |                          | GCTGAGCTGG<br>CGACTCGACC |                          |                          |                          |
| 2201 |                          | TGGGGTGTAC<br>ACCCCACATG |                          |                          |                          |
| 2251 | ****                     | AGGGCCAGGG<br>TCCCGGTCCC |                          |                          |                          |
| 2301 |                          | CTGAAGACTG<br>GACTTCTGAC |                          |                          |                          |
| 2351 |                          | GAAGCAGCTG<br>CTTCGTCGAC |                          |                          |                          |
| 2401 |                          | TCTGGGGCAA<br>AGACCCCGTT |                          |                          |                          |
| 2451 |                          | GAGACCTGGT<br>CTCTGGACCA |                          |                          |                          |
| 2501 |                          | GTTTGTGAAC<br>CAAACACTTG |                          |                          |                          |
| 2551 | CTGGAGAAGG<br>GACCTCTTCC | AGCCCATTGT<br>TCGGGTAACA | GGGGGCTGAG<br>CCCCCGACTC | ACCTTCTATG<br>TGGAAGATAC | TGGCTGGGGC<br>ACCGACCCCG |
| 2601 | TGCCAACAGG<br>ACGGTTGTCC | GAGACCAAGC<br>CTCTGGTTCG |                          |                          |                          |
| 2651 | GCAGGCAGAA<br>CGTCCGTCTT |                          |                          |                          | GAAGACTGCC<br>CTTCTGACGG |
| 2701 |                          |                          |                          |                          | AGGTGAACAT<br>TCCACTTGTA |
| 2751 |                          |                          |                          |                          | CAGCCTGATC<br>GTCGGACTAG |



| 2851 | GAGAAGGTGT<br>CTCTTCCACA | ACCTGGCCTG<br>TGGACCGGAC |            |            |                          |
|------|--------------------------|--------------------------|------------|------------|--------------------------|
| 2901 |                          | GACAAGCTGG<br>CTGTTCGACC |            | •          |                          |
| 2951 |                          | TGACAAGGCC<br>ACTGTTCCGG |            | -          |                          |
| 3001 |                          | TGGCCTCTGA<br>ACCGGAGACT |            |            |                          |
| 3051 | GATTGTGGCC<br>CTAACACCGG | TCCTGTGACA<br>AGGACACTGT |            |            |                          |
| 3101 |                          | CTGCTCCCCT<br>GACGAGGGGA |            |            |                          |
| 3151 |                          | TGATCCTGGT<br>ACTAGGACCA |            |            |                          |
| 3201 |                          | ATCCCTGCTG<br>TAGGGACGAC |            |            | · · · · · ·              |
| 3251 |                          | TGGCAGGTGG<br>ACCGTCCACC |            |            |                          |
| 3301 | AGGTTGAAGT               |                          | TCACTCCCGA | CGGACGACCA | CCCGACCGTA               |
| 3351 |                          | TTTGGCATCC<br>AAACCGTAGG |            |            |                          |
| 3401 | CCTCCATGAA<br>GGAGGTACTT | CAAGGAGCTG<br>GTTCCTCGAC |            |            | •                        |
|      | GCTGAGCACC<br>CGACTCGTGG | ACTTCTGTCG               | ACACGTCTAC | CGACACAAGT | AGGTGTTGAA               |
| 3501 |                          | GGGGGCATCG<br>CCCCCGTAGC |            |            |                          |
|      |                          | GTGTCTGTAG               | GTCTGGTTCC | TCGAGGTCTT | CGTCTAGTGG               |
|      |                          | TGAAGTCCCA               | CATGATGTCC | CTGAGGTCCT | TGGGGGACAC               |
| 3651 | GAAGGGCCCT<br>CTTCCCGGGA |                          |            |            | GTGGTGATCC<br>CACCACTAGG |
| 3701 | AGGACAACTC<br>TCCTGTTGAG |                          |            |            | CAAGATCATC<br>GTTCTAGTAG |

7 jure 26 D

| 3801  | GGATGAGGAC<br>CCTACTCCTG |                          | GCAGATCTGC<br>CGTCTAGACG |                          |                          |
|-------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| 3851  |                          |                          | CCCGTGCCTT<br>GGGCACGGAA |                          |                          |
| 3901. | ACTCCCACTG<br>TGAGGGTGAC |                          | ATAAAATGAG<br>TATTTTACTC |                          |                          |
| 3951  |                          |                          | TGGGGGGTGG<br>ACCCCCCACC |                          |                          |
| 4001  |                          |                          | AGCAGGCATG<br>TCGTCCGTAC |                          |                          |
| 4051  |                          |                          | ACTGAAATGT<br>TGACTTTACA |                          |                          |
| 4101  |                          |                          | GGGTCTTATG<br>CCCAGAATAC |                          |                          |
| 4151  |                          |                          | GCACCAACTC<br>CGTGGTTGAG |                          |                          |
| 4201  |                          |                          | ATGCCCCCAT<br>TACGGGGGTA |                          |                          |
| 4251  |                          |                          | TGGTCGCCCC<br>ACCAGCGGGG |                          |                          |
| 4301  |                          |                          | TGTCTGGAAC<br>ACAGACCTTG |                          |                          |
| 4351  |                          |                          | GCAGCCACCG<br>CGTCGGTGGC |                          |                          |
| 4401  |                          |                          | TGCAAACAGT<br>ACGTTTGTCA |                          |                          |
| 4451  |                          |                          | CTCTTTTGGC<br>GAGAAAACCG |                          | TCTTTGACCC<br>AGAAACTGGG |
| 4501  | GGGAACTTAA<br>CCCTTGAATT | TGTCGTTTCT<br>ACAGCAAAGA | CAGCAGCTGT<br>GTCGTCGACA | TGGATCTGCG<br>ACCTAGACGC | CCAGCAGGTT<br>GGTCGTCCAA |
| 4551  | TCTGCCCTGA<br>AGACGGGACT | AGGCTTCCTC<br>TCCGAAGGAG | CCCTCCCAAT               | GCGGTTTAAA<br>CGCCAAATTT | ACATAAATAA<br>TGTATTTATT |
| 4601  |                          |                          |                          |                          | TGCTGTCTTT<br>ACGACAGAAA |
| 4651  |                          |                          |                          |                          | GTCTCGGTCG<br>CAGAGCCAGC |

Figure 26E

| 4751 |                          |                     |   | GGGGTGGAGG<br>CCCCACCTCC |                          |
|------|--------------------------|---------------------|---|--------------------------|--------------------------|
| 4801 |                          | • • • • • • • • • • |   | AGATGATCCA<br>TCTACTAGGT |                          |
| 4851 |                          |                     |   | TTCAGTAGCA<br>AAGTCATCGT |                          |
| 4901 |                          |                     |   | AAAGCGGTTA<br>TTTCGCCAAT |                          |
| 4951 |                          |                     |   | TGGACTGTAT<br>ACCTGACATA |                          |
| 5001 |                          |                     |   | TTCATGTTGT<br>AAGTACAACA |                          |
| 5051 |                          |                     |   | TTTGTCATGT<br>AAACAGTACA |                          |
| 5101 | -                        |                     | • | TGTGACCTCC<br>ACACTGGAGG |                          |
| 5151 |                          |                     |   | CCACGGGCGG<br>GGTGCCCGCC |                          |
| 5201 |                          |                     |   | GTTGTGTTCC<br>CAACACAAGG |                          |
| 5251 |                          |                     |   | GGAGGGTGCC<br>CCTCCCACGG |                          |
| 5301 |                          |                     |   | TTACCCTCAC<br>AATGGGAGTG |                          |
| 5351 |                          |                     |   | CATGTCTACC<br>GTACAGATGG |                          |
| 5401 |                          |                     |   | TCAGCTGGGA<br>AGTCGACCCT |                          |
| 5451 | TTCCTGAGCA<br>AAGGACTCGT |                     |   |                          | AAATCACACC<br>TTTAGTGTGG |
| 5501 |                          |                     |   |                          | CCGTCATCCC<br>GGCAGTAGGG |
| 5551 | TGAGCAGGGG<br>ACTCGTCCCC |                     |   |                          | CATGTTTTCC<br>GTACAAAAGG |
| 5601 |                          |                     |   |                          | GCAGTTCTTG<br>CGTCAAGAAC |



| 5701 |                          | TTGACCAAGC<br>AACTGGTTCG |                          |                          |                          |
|------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| 5751 |                          | CATCTCGATC<br>GTAGAGCTAG |                          |                          |                          |
| 5801 |                          | TGTACGGCAG<br>ACATGCCGTC |                          |                          |                          |
| 5851 |                          | CACGGGCGCA<br>GTGCCCGCGT |                          |                          |                          |
| 5901 |                          | CGCTCCGGGC               |                          |                          |                          |
| 5951 |                          | TGCTGAAGCG<br>ACGACTTCGC |                          |                          |                          |
| 6001 |                          | ACCATGGTGT<br>TGGTACCACA |                          |                          |                          |
| 6051 |                          | CTTGCCCTTG<br>GAACGGGAAC |                          |                          |                          |
| 6101 |                          | CGTAGAGCTT<br>GCATCTCGAA |                          |                          |                          |
| 6151 |                          | CCGCAGGCCC<br>GGCGTCCGGG |                          |                          |                          |
| 6201 | TGAGCTCTGG<br>ACTCGAGACC | CCGTTCGGGG<br>GGCAAGCCCC | TCAAAAACCA<br>AGTTTTTGGT | GGTTTCCCCC<br>CCAAAGGGGG | ATGCTTTTTG<br>TACGAAAAAC |
| 6251 |                          | TACCTCTGGT<br>ATGGAGACCA |                          |                          |                          |
| 6301 | GAAAAGGCTG<br>CTTTTCCGAC | TCCGTGTCCC<br>AGGCACAGGG | CGTATACAGA<br>GCATATGTCT | CTTGAGAGGC<br>GAACTCTCCG | CTGTCCTCGA<br>GACAGGAGCT |
| 6351 | GCGGTGTTCC<br>CGCCACAAGG | GCGGTCCTCC<br>CGCCAGGAGG | TCGTATAGAA<br>AGCATATCTT | ACTCGGACCA<br>TGAGCCTGGT | CTCTGAGACA<br>GAGACTCTGT |
| 6401 | AAGGCTCGCG<br>TTCCGAGCGC | TCCAGGCCAG<br>AGGTCCGGTC | CACGAAGGAG<br>GTGCTTCCTC | GCTAAGTGGG<br>CGATTCACCC | AGGGGTAGCG<br>TCCCCATCGC |
| 6451 | GTCGTTGTCC<br>CAGCAACAGG | ACTAGGGGGT<br>TGATCCCCCA | CCACTCGCTC<br>GGTGAGCGAG | CAGGGTGTGA<br>GTCCCACACT | AGACACATGT<br>TCTGTGTACA |
| 6501 | CGCCCTCTTC<br>GCGGGAGAAG | GGCATCAAGG<br>CCGTAGTTCC | AAGGTGATTG<br>TTCCACTAAC | GTTTGTAGGT<br>CAAACATCCA | GTAGGCCACG<br>CATCCGGTGC |
| 6551 | TGACCGGGTG<br>ACTGGCCCAC | TTCCTGAAGG<br>AAGGACTTCC | GGGGCTATAA<br>CCCCGATATT | AAGGGGGTGG<br>TTCCCCCACC | GGGCGCGTTC<br>CCCGCGCAAG |

Figure 266

| 6651 | AGTACTCCCT<br>TCATGAGGGA |            | GGCATGACTT<br>CCGTACTGAA |            |                          |
|------|--------------------------|------------|--------------------------|------------|--------------------------|
| 6701 |                          |            | GATATTCACC<br>CTATAAGTGG |            |                          |
| 6751 |                          |            | GGTCAGAAAA<br>CCAGTCTTTT |            |                          |
| 6801 |                          |            | TAGAGGGCGT<br>ATCTCCCGCA |            | _                        |
| 6851 | CTCGCGTCCC               | AAACCAAAAA | GTCGCGATCG<br>CAGCGCTAGC | CGCGCGAGGA | ACCGGCGCTA               |
| 6901 | CAAATCGACG               | TGCATAAGCG | GCGCAACGCA<br>CGCGTTGCGT | GGCGGTAAGC | CCTTTCTGCC               |
| 6951 | ACCACGCGAG               | CAGCCCGTGG | AGGTGCACGC<br>TCCACGTGCG | CGGTTGGCGC | CAACACGTCC               |
| 7001 | CACTGTTCCA               | GTTGCGACCA | GGCTACCTCT<br>CCGATGGAGA | GGCGCATCCG | CGAGCAACCA               |
| 7051 | GGTCGTCTCC               | GCCGGCGGGA | TGCGCGAGCA<br>ACGCGCTCGT | CTTACCGCCA | TCCCCCAGAT               |
| 7101 | CGACGCAGAG               | CAGGCCCCCC | TCTGCGTCCA<br>AGACGCAGGT | GCCATTTCTG | GGGCCCGTCG               |
| 7151 | TCCGCGCGCA               | GCTTCATCAG | TATCTTGCAT<br>ATAGAACGTA | GGAACGTTCA | GATCGCGGAC               |
| 7201 | GACGGTACGC               | GCCCGCCGTT | GCGCGCGCTC<br>CGCGCGCGAG | CATACCCAAC | TCACCCCCTG               |
| 7251 | GGGTACCGTA               | CCCCACCCAC | AGCGCGGAGG<br>TCGCGCCTCC | GCATGTACGG | CGTTTACAGC               |
| 7301 | ATTTGCATCT               | CCCCGAGAGA | CTCATAAGGT               | TCTATACATC | GGTAGCATCT<br>CCATCGTAGA |
|      | AGGTGGCGCC               | TACGACCGCG | CGTGCATTAG               | CATATCAAGC | TGCGAGGGAG<br>ACGCTCCCTC |
|      | GCTCCTCCAG               | CCCTGGCTCC | AACGATGCCC               | GCCCGACGAG | TGCTCGGAAG<br>ACGAGCCTTC |
|      | TGATAGACGG               | ACTTCTACCG | TACACTCAAC               | CTACTATACC | TTGGACGCTG<br>AACCTGCGAC |
| 7501 |                          |            |                          |            | CGCACGAAGG               |

Figure 26 H

| 7601 |                          | - | <br>ATGATGTCAT<br>TACTACAGTA |                          |
|------|--------------------------|---|------------------------------|--------------------------|
| 7651 |                          |   | <br>GACAAACTCT<br>CTGTTTGAGA |                          |
| 7701 |                          |   | CCTCCGAACG<br>GGAGGCTTGC     |                          |
| 7751 |                          |   | GCGCAGCATC<br>CGCGTCGTAG     |                          |
| 7801 |                          |   | <br>GAGCGAGGTG<br>CTCGCTCCAC |                          |
| 7851 | CAAAGGTGTC<br>GTTTCCACAG |   | ACTGGTATTT<br>TGACCATAAA     |                          |
| 7901 |                          |   | AAGTCCGTGC<br>TTCAGGCACG     |                          |
| 7951 |                          |   | <br>GTTGAAGAGT<br>CAACTTCTCA |                          |
| 8001 |                          |   | AGGGTCCCGG<br>TCCCAGGGCC     |                          |
| 8051 |                          |   | ATCTCGTCAA<br>TAGAGCAGTT     |                          |
| 8101 |                          |   | GCGCGGGATG<br>CGCGCCCTAC     |                          |
| 8151 | •                        |   | GCTCTTCAGG<br>CGAGAAGTCC     |                          |
| 8201 |                          |   | TGAGGGTTGG<br>ACTCCCAACC     |                          |
| 8251 |                          |   | TTGCAGGTGG<br>AACGTCCACC     |                          |
| 8301 | TCCTAAACTG<br>AGGATTTGAC |   | CTGGGGTGAT<br>GACCCCACTA     |                          |
| 8351 | GTAAGCGGGT<br>CATTCGCCCA |   | CCAAGGTTCG<br>GGTTCCAAGC     |                          |
| 8401 | TCGCGCGGCA<br>AGCGCGCCGT |   | GCCGAACTTC<br>CGGCTTGAAG     |                          |
| 8451 |                          |   |                              | ATAGGTCTCT<br>TATCCAGAGA |

Figure 26I

| 8551 |            |            |            | GGAGTGGCTA<br>CCTCACCGAT |            |
|------|------------|------------|------------|--------------------------|------------|
|      |            |            |            |                          |            |
| 8601 | GAAAGTAGAA | GTCCCTGCGA | CGGGCCGAAC | ACTCGTGCTG               | GCTTTTGTAA |
|      | CTTTCATCTT | CAGGGACGCT | GCCCGGCTTG | TGAGCACGAC               | CGAAAACATT |
| 8651 | AAACGTGCGC | AGTACTGGCA | GCGGTGCACG | GGCTGTACAT               | CCTGCACGAG |
|      | TTTGCACGCG | TCATGACCGT | CGCCACGTGC | CCGACATGTA               | GGACGTGCTC |
| 8701 |            |            |            | GAGTGGGAAT               |            |
|      | CAACTGGACT | GCTGGCGCGT | GTTCCTTCGT | CTCACCCTTA               | AACTCGGGGA |
| 8751 |            |            |            | CTTCGGCTGC               |            |
|      | GCGGACCGCC | CAAACCGACC | ACCAGAAGAT | GAAGCCGACG               | AACAGGAACT |
| 8801 |            |            |            | GATCGGACCA               |            |
|      | GGCAGACCGA | CGAGCTCCCC | TCAATGCCAC | CTAGCCTGGT               | CCTCCCCCCC |
| 8851 |            |            |            | CGGTCGGAGC               |            |
|      | GCTCGGGTTT | CAGGTCTACA | GGCGCGCCC  | GCCAGCCTCG               | AACTACTGTT |
| 8901 | CATCGCGCAG | ATGGGAGCTG | TCCATGGTCT | GGAGCTCCCG               | CGGCGTCAGG |
|      |            |            |            | CCTCGAGGGC               |            |
| 8951 | TCAGGCGGGA | GCTCCTGCAG | GTTTACCTCG | CATAGACGGG               | TCAGGGCGCG |
|      | AGTCCGCCCT | CGAGGACGTC | CAAATGGAGC | GTATCTGCCC               | AGTCCCGCGC |
| 9001 | GGCTAGATCC | AGGTGATACC | TAATTTCCAG | GGGCTGGTTG               | GTGGCGGCGT |
|      | CCGATCTAGG | TCCACTATGG | ATTAAAGGTC | CCCGACCAAC               | CACCGCCGCA |
| 9051 |            |            |            | GCGCGACTAC               |            |
|      | GCTACCGAAC | GTTCTCCGGC | GTAGGGGCGC | CGCGCTGATG               | CCATGGCGCG |
| 9101 |            |            |            | GATGATGCAT               |            |
|      | CCGCCCGCCA | CCCGGCGCCC | CCACAGGAAC | CTACTACGTA               | GATTTTCGCC |
| 9151 |            |            |            | GGCTCCGGAC               |            |
|      | ACTGCGCCCG | CTCGGGGGCC | TCCATCCCCC | CCGAGGCCTG               | GGCGGCCCTC |
| 9201 |            |            |            |                          | TGGTGCTGCG |
| •    | TCCCCCGTCC | CCGTGCAGCC | ececcece   | CCCGTCCTCG               | ACCACGACGC |
| 9251 | CGCGTAGGTT | GCTGGCGAAC | GCGACGACGC | GGCGGTTGAT               | CTCCTGAATC |
|      | GCGCATCCAA | CGACCGCTTG | CGCTGCTGCG | CCGCCAACTA               | GAGGACTTAG |
| 9301 | TGGCGCCTCT | GCGTGAAGAC | GACGGGCCCG | GTGAGCTTGA               | ACCTGAAAGA |
|      | ACCGCGGAGA | CGCACTTCTG | CTGCCCGGGC | CACTCGAACT               | TGGACTTTCT |
| 9351 | GAGTTCGACA | GAATCAATTT | CGGTGTCGTT | GACGGCGGCC               | TGGCGCAAAA |
|      | CTCAAGCTGT | CTTAGTTAAA | GCCACAGCAA | CTCCCCCCG                | ACCGCGTTTT |
| 9401 | TCTCCTGCAC | GTCTCCTGAG | TTGTCTTGAT | AGGCGATCTC               | GGCCATGAAC |
|      | AGAGGACGTG | CAGAGGACTC | AACAGAACTA | TCCGCTAGAG               | CCGGTACTTG |

Figure 26. J 64/144

| 9501  |                          | TCGTTGGAAA<br>AGCAACCTTT |                          |                          |                          |
|-------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| 9551  |                          | GTTCCAGACG<br>CAAGGTCTGC |                          |                          |                          |
| 9,601 |                          | TGACCACCTG<br>ACTGGTGGAC |                          |                          |                          |
| 9651  |                          | TTTCGCAGGC<br>AAAGCGTCCG |                          |                          |                          |
| 9701  |                          | CACGAAGAAG<br>GTGCTTCTTC |                          |                          |                          |
| 9751  |                          | CCAAGGCCTC<br>GGTTCCGGAG |                          |                          |                          |
| 9801  |                          | AAAAACTGGG<br>TTTTTGACCC |                          |                          |                          |
| 9851  |                          | GATGAGCTCG<br>CTACTCGAGC |                          |                          |                          |
| 9901  |                          | CCTCTTCTTC<br>GGAGAAGAAG |                          |                          |                          |
| 9951  |                          | TCTTCTGGCG<br>AGAAGACCGC |                          |                          |                          |
| 10001 |                          | CGGGAGGCGG<br>GCCCTCCGCC |                          |                          |                          |
| 10051 |                          | TGGTCTCGGT<br>ACCAGAGCCA |                          |                          |                          |
| 10101 |                          | CCGCCCGTCA<br>GGCGGGCAGT |                          |                          |                          |
| 10151 |                          |                          |                          |                          | TTGTTGTGTA<br>AACAACACAT |
| 10201 | GGTACTCCGC<br>CCATGAGGCG | CGCCGAGGGA<br>GCGGCTCCCT | CCTGAGCGAG<br>GGACTCGCTC | TCCGCATCGA<br>AGGCGTAGCT | CCGGATCGGA<br>GGCCTAGCCT |
| 10251 |                          |                          |                          |                          | GGTAGGCTGA<br>CCATCCGACT |
| 10301 |                          |                          |                          |                          | TCTGGCGGAG<br>AGACCGCCTC |
| 10351 |                          |                          |                          |                          | GGCGGATGGT<br>CCGCCTACCA |

Figure 26 K

| 10451 | CGGCCATGCC<br>GCCGGTACGG |       | GGCGCAGGTC<br>CCGCGTCCAG  |                          |
|-------|--------------------------|-------|---------------------------|--------------------------|
| 10501 | • •                      | <br>  | TCTTCTCCTT<br>AGAAGAGGAA  |                          |
| 10551 |                          | <br>  | GGCGGAGTTT<br>CCGCCTCAAA  |                          |
| 10601 |                          |       | CGAAGCCCCT<br>GCTTCGGGGA  | •                        |
| 10651 |                          |       | GCTAATATGG<br>CGATTATACC  |                          |
|       | CTGCGTGAGG<br>GACGCACTCC | <br>  |                           |                          |
| 10751 |                          | <br>  | CCATAACGGA<br>GGTATTGCCT  |                          |
| 10801 |                          | <br>  | TACCTGAGAC<br>ATGGACTCTG  |                          |
| 10851 |                          | <br>  | CCGCACCAGG<br>GGCGTGGTCC  |                          |
| 10901 |                          |       | AGAGGGGCCA<br>TCTCCCCGGT  |                          |
| 10951 |                          | <br>- | ATAAGGCGAT<br>TATTCCGCTA  |                          |
| 11001 |                          | <br>  | GGCGGTGGTG<br>CCGCCACCAC  |                          |
| 11051 |                          | <br>  | GCAGCGGCAA<br>CGTCGCCGTT  |                          |
| 11101 |                          | <br>  | GCGCAATCGT<br>CGCGTTAGCA  |                          |
| 11151 |                          | <br>  | CACTCTTCCG<br>GTGAGAAGGC  | TGGTCTGGTG<br>ACCAGACCAC |
| 11201 |                          |       |                           | GAGCCCCGTA<br>CTCGGGGCAT |
| 11251 | TCCGGCCGTC<br>AGGCCGGCAG |       | CCGCCCGCGT<br>GGCGGGGCGCA |                          |
| 11301 |                          |       |                           | TTCCTTCCAG<br>AAGGAAGGTC |

Figure 26L

| 11401 |                          |                          |                          | AGTGGCTCGC<br>TCACCGAGCG |            |
|-------|--------------------------|--------------------------|--------------------------|--------------------------|------------|
| 11451 |                          |                          |                          | GGGACCCCCG<br>CCCTGGGGGC |            |
| 11501 |                          |                          |                          | TTGCCTCCCC<br>AACGGAGGGG |            |
| 11551 |                          |                          |                          | GACGAGCCCC<br>CTGCTCGGGG |            |
| 11601 | AAGGGTCTAC               | GTAGGCCACG               | ACGCCGTCTA               | GCGCCCCCT                | GGAGTCGTCG |
| 11651 | CCGTTCTCGT               | TCTCGTCGCC               | GTCTGTACGT               | GGGCACCCTC<br>CCCGTGGGAG | GGGAGGAGGA |
| 11701 | TGGCGCAGTC               | CTCCCCGCTG               | TAGGCGCCAA               | GACGCGGCAG<br>CTGCGCCGTC | GTCTACCACT |
| 11751 | AATGCTTGGG               | GGCGCCGCGG               | CCCGGGCCGT               | CTACCTGGAC<br>GATGGACCTG | AACCTCCTCC |
| 11801 | CGCTCCCGGA               | CCGCGCCGAT               | CCTCGCGGGA               | CTCCTGAGCG<br>GAGGACTCGC | CGTGGGTTCC |
| 11851 | CACGTCGACT               | TCGCACTATG               | CGCACTCCGC               | TACGTGCCGC<br>ATGCACGGCG | CCGTCTTGGA |
| 11901 | CAAAGCGCTG               | GCGCTCCCTC               | TCCTCGGGCT               | GGAGATGCGG<br>CCTCTACGCC | CTAGCTTTCA |
| 11951 | AGGTGCGTCC               | CGCGCTCGAC               | GCCGTACCGG               | TGAATCGCGA<br>ACTTAGCGCT | CGCCAACGAC |
| 12001 | GCGCTCCTCC               | TGAAACTCGG               | GCTGCGCGCT               | ACCGGGATTA<br>TGGCCCTAAT | CAGGGCGCGC |
| 12051 | GCGTGTGCAC               | CGCCGGCGGC               | TGGACCATTG               | CGCATACGAG               | GTCTGCCACT |
|       | _                        | ATTGAAAGTT               | TTTTCGAAAT               | TGTTGGTGCA               | CGCATGCGAA |
|       |                          | TCCTCCACCG               | ATATCCTGAC               | TACGTAGACA               | CCCTGAAACA |
|       |                          | CTCGTTTTGG               | GTTTATCGTT               | CGGCGAGTAC               | CGCGTCGACA |
| 12251 | TCCTTATAGT<br>AGGAATATCA | GCAGCACAGC<br>CGTCGTGTCG | AGGGACAACG<br>TCCCTGTTGC | AGGCATTCAG<br>TCCGTAAGTC | CCTACGCGAC |

7 igure 26 M

| 12351 |                          |                          |                          | CTTGAGCCTG<br>GAACTCGGAC |                          |
|-------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| 12401 |                          |                          |                          | TGGGCAAGTT<br>ACCCGTTCAA |                          |
| 12451 |                          |                          |                          | GACAAGGAGG<br>CTGTTCCTCC |                          |
| 12501 |                          |                          |                          | GCTTACCTTG<br>CGAATGGAAC |                          |
| 12551 |                          |                          |                          | AGGCCGTGAG<br>TCCGGCACTC |                          |
| 12601 |                          |                          |                          | CACAGCCTGC<br>GTGTCGGACG |                          |
| 12651 |                          |                          |                          | CGAGTCCTAC<br>GCTCAGGATG |                          |
| 12701 |                          |                          |                          | GCGCCCTGGA<br>CGCGGGACCT |                          |
| 12751 |                          |                          |                          | CGCGCTGGCA<br>GCGCGACCGT |                          |
| 12801 |                          |                          |                          | CGAGCCAGAG<br>GCTCGGTCTC |                          |
| 12851 |                          |                          |                          | GCAAGACGCA<br>CGTTCTGCGT |                          |
| 12901 |                          |                          |                          | CCGGCCTTAA<br>GGCCGGAATT |                          |
| 12951 |                          |                          |                          | TCGCTGACTG<br>AGCGACTGAC |                          |
| 13001 | TGACGCGTTC<br>ACTGCGCAAG | CGGCAGCAGC<br>GCCGTCGTCG | CGCAGGCCAA<br>GCGTCCGGTT | CCGGCTCTCC<br>GGCCGAGAGG | GCAATTCTGG<br>CGTTAAGACC |
| 13051 |                          |                          |                          |                          | GGTGCTGGCG<br>CCACGACCGC |
| 13101 |                          |                          |                          |                          | ACGAGGCCGG<br>TGCTCCGGCC |
| 13151 |                          |                          |                          |                          | AACAGCGGCA<br>TTGTCGCCGT |
| 13201 |                          |                          |                          |                          | CGAGGCCGTG<br>GCTCCGGCAC |

Figure 26 N

| 13301 |                          | TTCCTGAGTA<br>AAGGACTCAT |   |                          |                          |
|-------|--------------------------|--------------------------|---|--------------------------|--------------------------|
| 13351 |                          | CAACTTTGTG<br>GTTGAAACAC |   |                          |                          |
| 13401 |                          | AGGTGTACCA<br>TCCACATGGT |   |                          |                          |
| 13451 |                          | CTGCAGACCG<br>GACGTCTGGC | · · · · · ·                             |                          |                          |
| 13501 |                          | GGGGGTGCGG               |   |                          |                          |
| 13551 |                          | CGCCCAACTC<br>GCGGGTTGAG |   |                          |                          |
| 13601 |                          | GGCAGCGTGT<br>CCGTCGCACA | • |                          | CACTTGCTGA<br>GTGAACGACT |
| 13651 |                          | CGAGGCCATA<br>GCTCCGGTAT |   | · - · - <del>-</del> · · |                          |
| 13701 |                          | CAAGTGTCAG<br>GTTCACAGTC |   |                          |                          |
| 13751 |                          | ACCCTAAACT<br>TGGGATTTGA |   |                          |                          |
| 13801 |                          | CAGTTTAAAC<br>GTCAAATTTG |   |                          |                          |
| 13851 |                          | TGAGCCTTAA<br>ACTCGGAATT |   |                          | =                        |
| 13901 |                          | ATGACCGCGC<br>TACTGGCGCG |   |                          |                          |
| 13951 |                          | TATCAACCGC<br>ATAGTTGGCG |   |                          |                          |
| 14001 | GTGAACCCCG<br>CACTTGGGGC | AGTATTTCAC<br>TCATAAAGTG |   |                          |                          |
| 14051 | GCCCCCTGGT<br>CGGGGGACCA | TTCTACACCG<br>AAGATGTGGC |   |                          |                          |
| 14101 | GATTCCTCTG<br>CTAAGGAGAC | GGACGACATA<br>CCTGCTGTAT |   |                          |                          |
| 14151 | ACCCTGCTAG<br>TGGGACGATC | AGTTGCAACA<br>TCAACGTTGT |   |                          |                          |

7, gure 260

| 14251 |            |            |            | GCTTGATAGG<br>CGAACTATCC |                          |
|-------|------------|------------|------------|--------------------------|--------------------------|
| 14301 |            |            |            | GGCGAGGAGG<br>CCGCTCCTCC |                          |
| 14351 |            |            |            | AAACCTGCCT<br>TTTGGACGGA |                          |
| 14401 |            |            |            | AGATGAGTAG<br>TCTACTCATC |                          |
| 14451 |            |            |            | GCGCGCCGC                |                          |
| 14501 |            |            |            | GTGGGAGGAC<br>CACCCTCCTG |                          |
| 14551 | GTCTGCTGTC | GTCGCAGGAC | CTAAACCCTC | GGAGTGGCAA<br>CCTCACCGTT | GGGCAAACGC               |
| 14601 |            |            |            | TAAAAAAAAA<br>ATTTTTTTT  |                          |
| 14651 | ACGTTTTATT | TTTTGAGTGG | TTCCGGTACC | CACCGAGCGT<br>GTGGCTCGCA | ACCAAAAGAA               |
| 14701 | CATAAGGGGA | ATCATACGCC | GCGCGCCGCT | TGTATGAGGA<br>ACATACTCCT | TCCAGGAGGA               |
| 14751 |            |            |            | CCAGTGGCGG<br>GGTCACCGCC |                          |
| 14801 | AAGAGGGAAG | CTACGAGGGG | ACCTGGGCGG | GTTTGTGCCT<br>CAAACACGGA | GGCGCCATGG               |
| 14851 | ÄCGCCGGATG | GCCCCCCTCT | TTGTCGTAGG | GTTACTCTGA<br>CAATGAGACT | CAACCGTGGG               |
| 14901 | GATAAGCTGT | GGTGGGCACA | CATGGACCAC | CTGTTGTTCA               | CAACGGATGT<br>GTTGCCTACA |
|       | CCGTAGGGAC | TTGATGGTCT | TGCTGGTGTC | GTTGAAAGAC               | ACCACGGTCA<br>TGGTGCCAGT |
|       | AAGTTTTGTT | ACTGATGTCG | GGCCCCCTCC | GTTCGTGTGT               | GACCATCAAT<br>CTGGTAGTTA |
| •     | GAACTGCTGG | CCAGCGTGAC | CCCGCCGCTG | GACTTTTGGT               | TCCTGCATAC<br>AGGACGTATG |
| 15101 |            |            |            |                          | TTTAAGGCGC<br>AAATTCCGCG |

Figure 26 P

| WO 02/022080 |                          |  |                          | PCT/US01/28861   |
|--------------|--------------------------|--|--------------------------|--|
| 15151        | GGGTGATGGT<br>CCCACTACCA |  | ACAATCAGGT<br>TGTTAGTCCA |  |
|              | TACGAGTGGG<br>ATGCTCACCC |  |                          |  |
| 15251        | GACCATAGAC<br>CTGGTATCTG |  | GGAGCACTAC<br>CCTCGTGATG |  |
| 15301        | GCAGACAGAA<br>CGTCTGTCTT |  | TCGGGGTAAA<br>AGCCCCATTT |  |
| 15351        | CGCAACTTCA<br>GCGTTGAAGT |  | ACTGGTCTTG<br>TGACCAGAAC |  |
| 15401        | GGTATATACA<br>CCATATATGT |  | CATCATTTTG<br>GTAGTAAAAC |  |
| 15451        | GCGGGGTGGA<br>CGCCCCACCT |  | GCAACTTGTT<br>CGTTGAACAA |  |
| 15501        | AAGCGGCAAC<br>TTCGCCGTTG |  | ATCACCTACG<br>TAGTGGATGC |  |
| 15551        | GGGTGGTAAC<br>CCCACCATTG |  | GGAÇGCCTAC<br>CCTGCGGATG |  |
| 15601        | TGAAAGATGA<br>ACTTTCTAÇT |  | GCGCAGGCGG<br>CGCGTCCGCC |  |
| 15651        | AGTGGCAGCG<br>TCACCGTCGC |  | GCGGCAGCCG<br>CGCCGTCGGC |  |
| 15701        | GCCGGTGGAG<br>CGGCCACCTC |  | TCGCGGCGAC<br>AGCGCCGCTG |  |
| 15751        |                          |  | AAGCAGCGGC<br>TTCGTCGCCG |  |
| 15801        | GCCCCCGCTG<br>CGGGGGCGAC |  | •                        | AACCGGTGAT<br>TTGGCCACTA   |
| 15851        | CAAACCCCTG<br>GTTTGGGGAC |  | CAGTTACAAC<br>GTCAATGTTG | A. Control of the Con |
| 15901        | ATGACAGCAC<br>TACTGTCGTG |  | GGTACCTTGC<br>CCATGGAACG |  |
| 15951        |                          |  |                          | GCACTCCTGA<br>CGTGAGGACT   |
| 16001        |                          |  | GTCGTTGCCA<br>CAGCAACGGT | GACATGATGC<br>CTGTACTACG   |
| 16051        |                          |  | AGATCAGCAA<br>TCTAGTCGTT | CTTTCCGGTG<br>GAAAGGCCAC   |

Figure 26 Q

| 16151 | <br> | <br>TACCTCTCTG<br>ATGGAGAGAC |                          |
|-------|------|------------------------------|--------------------------|
| 16201 |      | cececcecc<br>cececcecc       |                          |
| 16251 | <br> | <br>CTCACAGATC<br>GAGTGTCTAG |                          |
| 16301 |      | GCGAGTGACC<br>CGCTCACTGG     |                          |
| 16351 | <br> | <br>AGGCCCTGGG<br>TCCGGGACCC |                          |
| 16401 |      | GCAAGCATGT<br>CGTTCGTACA     |                          |
| 16451 |      | GCGCTTCCCA<br>CGCGAAGGGT     |                          |
| 16501 |      | ACCCAGTGCG<br>TGGGTCACGC     |                          |
| 16551 |      | CGCGGCCGCA<br>GCGCCGCGT      |                          |
| 16601 |      | GGAGGAGGCG<br>CCTCCTCCGC     |                          |
| 16651 |      | ACGCGGCCAT<br>TGCGCCGGTA     |                          |
| 16701 |      | AAGAGACGGC<br>TTCTCTGCCG     |                          |
| 16751 |      | TGCCGCCCAA<br>ACGGCGGGTT     |                          |
| 16801 |      | GCCGACGGGC<br>CGGCTGCCCG     | GGCCATGCGG<br>CCGGTACGCC |
| 16851 |      |                              | CCAGGTCCAG<br>GGTCCAGGTC |
| 16901 |      |                              | ATGACTCAGG<br>TACTGAGTCC |
| 16951 |      |                              | CGGCCTGCGC               |
| 17001 |      |                              | GAAAAAACTA<br>CTTTTTTGAT |

7igure 26 R

| 17101 |                          | -          | AAAGAAGAGA<br>TTTCTTCTCT |            |                          |
|-------|--------------------------|------------|--------------------------|------------|--------------------------|
| 17151 |                          |            | GAAGGAAGAG<br>CTTCCTTCTC |            |                          |
| 17201 |                          | _          | AAAAGAAAGA<br>TTTTCTTTCT |            |                          |
| 17251 |                          |            | GCTACCGCGC<br>CGATGGCGCG |            |                          |
| 17301 |                          |            | TGTTTTGCGA<br>ACAAAACGCT |            |                          |
| 17351 |                          |            | CCCGCACCTA<br>GGGCGTGGAT |            |                          |
| 17401 | ACATGCCGCT               | GCTCCTGGAC | CTTGAGCAGG<br>GAACTCGTCC | GGTTGCTCGC | GGAGCCCCTC               |
| 17451 |                          |            | TAAGGACATG<br>ATTCCTGTAC |            |                          |
| 17501 | ••••                     |            | TAAAGCCCGT<br>ATTTCGGGCA |            |                          |
| 17551 |                          |            | GAAAAGCGCG<br>CTTTTCGCGC |            |                          |
| 17601 | CTGAACCGTG               | GGTGGCACGT | GCTGATGGTA<br>CGACTACCAT | GGGTTCGCGG | TCGCTGACCT               |
| 17651 | TCTACAGAAC               | CTTTTTTACT | CCGTGGAACC<br>GGCACCTTGG | ACCCGACCTC | GGGCTCCAGG               |
| 17701 |                          |            | GTGGCGCCGG<br>CACCGCGGCC |            |                          |
| 17751 | CTGCAAGTCT               | ATGGGTGATG |                          | TCATAACGGT | GGCGGTGTCT               |
|       |                          | TGTGTTTGCA | GGGGCCAACG               | GAGTCGCCAC | CGCCTACGGC               |
|       |                          | CCAGCGACGC | CGGCGCAGGT               | TCTGGAGATG | CCTCCACGTT               |
| 17901 | ACGGACCCGT<br>TGCCTGGGCA |            |                          |            | CGCGCCGTTC               |
| 17951 |                          |            |                          |            | GCCCTACATC<br>CGGGATGTAG |

Figure 265

| 18051 |                          |      | ACTGGAACCC<br>TGACCTTGGG |                          |
|-------|--------------------------|------|--------------------------|--------------------------|
| 18101 | TCGCCGTCGC<br>AGCGGCAGCG |      | TTCCGTGCGC<br>AAGGCACGCG |                          |
| 18151 |                          |      | CAGCGCGCTA<br>GTCGCGCGAT |                          |
| 18201 |                          | <br> | GCAGATATGG<br>CGTCTATACC |                          |
| 18251 |                          |      | AGGAAGAATG<br>TCCTTCTTAC |                          |
| 18301 |                          |      | GCATGCGTCG<br>CGTACGCAGC |                          |
| 18351 |                          |      | CGCGGCGGTA<br>GCGCCGCCAT |                          |
| 18401 |                          |      | CGCCGTGCCC               |                          |
| 18451 | •••••                    | <br> | TAAAAACAAG<br>ATTTTTGTTC |                          |
| 18501 |                          |      | CGCTCGCTTG<br>GCGAGCGAAC |                          |
| 18551 |                          |      | CGTCTCTGGC<br>GCAGAGACCG |                          |
| 18601 |                          |      | GATATCGGCA<br>CTATAGCCGT |                          |
| 18651 |                          |      | GTGGAGCGGC<br>CACCTCGCCG |                          |
| 18701 |                          |      |                          | CAGCAGCACA<br>GTCGTCGTGT |
| 18751 |                          |      |                          | AACAAAAGGT<br>TTGTTTTCCA |
| 18801 | GGTAGATGGC<br>CCATCTACCG |      |                          | CTGGCCAACC<br>GACCGGTTGG |
| 18851 |                          |      |                          | CCCTCCCGTA<br>GGGAGGGCAT |
| 18901 |                          |      |                          | GGCGTGGCGA<br>CCGCACCGCT |

Figure 26.T

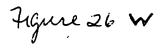
| 19001 | GTACGAGGAG<br>CATGCTCCTC |  |                          |
|-------|--------------------------|--|--------------------------|
| 19051 | CCATGGCTAC<br>GGTACCGATG |  |                          |
| 19101 | CCTCCCCCG<br>GGAGGGGGC   |  |                          |
| 19151 | CGTTGTTGTA<br>GCAACAACAT |  |                          |
| 19201 | GTCCGCGATC<br>CAGGCGCTAG |  |                          |
| 19251 | AACAGCATCG<br>TTGTCGTAGC |  |                          |
| 19301 | CTGATAGCTA<br>GACTATCGAT |  |                          |
| 19351 | CAGAGGAGCT<br>GTCTCCTCGA |  |                          |
| 19401 | TTCGATGATG<br>AAGCTACTAC |  |                          |
| 19451 | CGGAGTACCT<br>GCCTCATGGA |  |                          |
| 19501 | TACTTCAGCC<br>ATGAAGTCGG |  |                          |
| 19551 | CGACGTGACC<br>GCTGCACTGG |  |                          |
| 19601 | TGGACCGTGA<br>ACCTGGCACT |  |                          |
| 19651 | GTGGGTGATA<br>CACCCACTAT |  | TCCACGTACT<br>AGGTGCATGA |
| 19701 |                          |  | GCCCTACTCT CGGGATGAGA    |
| 19751 |                          |  | ATCCTTGCGA<br>TAGGAACGCT |
| 19801 |                          |  | GAAGAGGACG<br>CTTCTCCTGC |
| 19851 |                          |  | AAAAACTCAC<br>TTTTTGAGTG |

Figure 26 U

| 19951 | TCAAATAGGT<br>AGTTTATCCA                |   |                          |   |                          |
|-------|---|---|--------------------------|---|--------------------------|
| 20001 |   |   | GAATCTCAGT<br>CTTAGAGTCA |   |                          |
| 20051 | •••••                                   |   | AAAAAAGACT<br>TTTTTTCTGA |   |                          |
| 20101 |   |   | CAAATGAAAA<br>GTTTACTTTT |   |                          |
| 20151 |   |   | CTAGAAAGTC<br>GATCTTTCAG |   |                          |
| 20201 |   |   | AGGCAATGGT<br>TCCGTTACCA |   |                          |
| 20251 |   |   | TAGATATAGA<br>ATCTATATCT |   |                          |
| 20301 | • | * | GAAGGTAACT<br>CTTCCATTGA |   |                          |
| 20351 |   |   | TAATTACATT<br>ATTAATGTAA |   |                          |
| 20401 |   |   | GCACGGGTAA<br>CGTGCCCATT | , |                          |
| 20451 |   |   | GTTGTAGATT<br>CAACATCTAA |   |                          |
| 20501 |   |   | TGATTCCATT<br>ACTAAGGTAA |   |                          |
| 20551 |   |   | TTGACAGCTA<br>AACTGTCGAT | , |                          |
| 20601 |   |   | GATGAACTTC<br>CTACTTGAAG |   |                          |
| 20651 | GGÄGGTGTGA<br>CCTCCACACT                |   |                          |   | CTAAAACAGG<br>GATTTTGTCC |
|       | TCAGGAAAAT<br>AGTCCTTTTA                |   |                          |   | GATAAAAATG<br>CTATTTTTAC |
| 20751 |   |   |                          |   | AAATGCCAAC<br>TTTACGGTTG |
| 20801 |   |   |                          |   | TGCCCGACAA<br>ACGGGCTGTT |

Tigure 26 V

| 20901 | ACGACTACAT | GAACAAGCGA | GTGGTGGCTC | CCGGGCTAGT   | GGACTGCTAC |
|-------|------------|------------|------------|--------------|------------|
|       |            | CTTGTTCGCT | -          |              |            |
| 20951 |            | GAGCACGCTG |            |              |            |
|       |            | CTCGTGCGAC |            |              | •          |
| 21001 |            | CACCGCAATG |            | <del>-</del> |            |
|       |            | GTGGCGTTAC |            |              | -          |
| 21051 |            | CTATGTGCCC |            |              |            |
|       |            | GATACACGGG |            |              |            |
| 21101 |            | ACCTCCTTCT |            |              |            |
|       | CGGTAATTTT | TGGAGGAAGA | GGACGGCCCG | AGTATGTGGA   | TGCTCACCTT |
| 21151 |            | GATGTTAACA |            |              |            |
|       |            | CTACAATTGT |            |              |            |
| 21201 |            | CGGAGCCAGC | •          |              |            |
|       | ATTCCCAACT | GCCTCGGTCG | TAATTCAAAC | TATCGTAAAC   | GGAAATGCGG |
| 21251 |            | CCATGGCCCA |            |              |            |
|       | TGGAAGAAGG | GGTACCGGGT | GTTGTGGCGG | AGGTGCGAAC   | TCCGGTACGA |
| 21301 |            | ACCAACGACC |            |              |            |
|       |            | TGGTTGCTGG |            |              |            |
| 21351 |            | CCCTATACCC |            |              |            |
|       | TGTACGAGAT | GGGATATGGG | CGGTTGCGAT | GGTTGCACGG   | GTATAGGTAG |
| 21401 |            | ACTGGGCGGC |            |              |            |
|       |            | TGACCCGCCG |            |              |            |
| 21451 |            | ACCCCATCAC |            |              |            |
| •     |            | TGGGGTAGTG |            |              |            |
| 21501 |            | TATACCCTAC |            |              |            |
|       |            | ATATGGGATG |            |              |            |
| 21551 |            | TGGCCATTAC |            |              | •          |
|       | AAATTCTTCC | ACCGGTAATG | GAAACTGAGA | AGACAGTCGA   | CCGGACCGTT |
| 21601 | TGACCGCCTG | •          |            |              |            |
|       | ACTGGCGGAC | GAATGGGGGT | TGCTCAAACT | TTAATTCGCG   | AGTCAACTGC |
| 21651 | GGGAGGGTTA |            |            |              |            |
|       | CCCTCCCAAT | GTTGCAACGG | GTCACATTGT | ACTGGTTTCT   | GACCAAGGAC |
| 21701 | GTACAAATGC |            | · · ·      |              |            |
|       | CATGTTTACG | ATCGATTGAT | ATTGTAACCG | ATGGTCCCGA   | AGATATAGGG |
| 21751 | AGAGAGCTAC | AAGGACCGCA | TGTACTCCTT | CTTTAGAAAC   | TTCCAGCCCA |
|       | TCTCTCGATG | TTCCTGGCGT | ACATGAGGAA | GAAATCTTTG   | AAGGTCGGGT |



| 21851 |            | ACCAACACAA<br>TGGTTGTGTT |            |            |                          |
|-------|------------|--------------------------|------------|------------|--------------------------|
| 21901 |            | GAAGGACAGG<br>CTTCCTGTCC |            |            |                          |
| 21951 | ATCCGTTCTG | CGCAGTTGAC<br>GCGTCAACTG | TCGTAATGGG | TCTTTTTCAA | AGAAACGCTA               |
| 22001 | GCGTGGGAAA | GGCGCATCCC<br>CCGCGTAGGG | TAAGAGGTCA | TTGAAATACA | GGTACCCGCG               |
| 22051 | TGAGTGTCTG | CTGGGCCAAA<br>GACCCGGTTT | TGGAAGAGAT | GCGGTTGAGG | CGGGTGCGCG               |
| 22101 | ATCTGTACTG | TTTTGAGGTG<br>AAAACTCCAC | CTAGGGTACC | TGCTCGGGTG | GGAAGAAATA               |
| 22151 | CAAAACAAAC | AAGTCTTTGA<br>TTCAGAAACT | GCACCAGGCA | CACGTGGTCG | GCGTGGCGCC               |
| 22201 | GCAGTAGCTT | ACCGTGTACC<br>TGGCACATGG | ACGCGTGCGG | GAAGAGCCGG | CCGTTGCGGT               |
| 22251 | GTTGTATTTC | AAGCAAGCAA<br>TTCGTTCGTT | GTAGTTGTTG | TCGACGGCGG | TACCCGAGGT               |
| 22301 | CACTCGTCCT | ACTGAAAGCC<br>TGACTTTCGG | TAACAGTTTC | TAGAACCAAC | ACCCGGTATA               |
| 22351 | AAAAACCCGT | GGATACTGTT               | CGCGAAAGGT | CCGAAACAAA |                          |
| 22401 | CGAGCGGACG | GCCATAGTCA<br>CGGTATCAGT | TATGCCGGCC | AGCGCTCTGA | CCCCCGCATG               |
| 22451 | TGACCTACCG | GAAACGGACC               | TTGGGCGTGA | GTTTTTGTAC | CTACCTCTTT<br>GATGGAGAAA |
| 22501 | CTCGGGAAAC | CGAAAAGACT               | GGTCGCTGAG | TTCGTCCAAA | ACCAGTTTGA<br>TGGTCAAACT |
|       | CATGCTCAGT | GAGGACGCGG               | CATCGCGGTA | ACGAAGAAGG | CCCGACCGCT<br>GGGCTGGCGA |
|       | CATATTGCGA | CCTTTTCAGG               | TGGGTTTCGC | ATGTCCCCGG | CAACTCGGCC<br>GTTGAGCCGG |
|       | CGGACACCTG | ATAAGACGAC               | GTACAAAGAG | GTGCGGAAAC | CCAACTGGCC<br>GGTTGACCGG |
| 22701 |            |                          |            |            | ACCGGGGTAC<br>TGGCCCCATG |

Figure 26 X

| 22801 | CAGGAACAGC<br>GTCCTTGTCG | TCTACAGCTT<br>AGATGTCGAA                |                          |                          |                          |
|-------|--------------------------|---|--------------------------|--------------------------|--------------------------|
| 22851 | CCACAGTGCG<br>GGTGTCACGC | CAGATTAGGA<br>GTCTAATCCT                |                          |                          |                          |
| 22901 |                          | ATGTACTAGA<br>TACATGATCT                |                          |                          |                          |
| 22951 | TTGTACACTC<br>AACATGTGAG | TCGGGTGATT<br>AGCCCACTAA                | ATTTACCCCC<br>TAAATGGGGG | ACCCTTGCCG<br>TGGGAACGGC | TCTGCGCCGT<br>AGACGCGGCA |
| 23001 |                          | AAGGGGTTCT<br>TTCCCCAAGA                |                          |                          |                          |
| 23051 | ACACGTTGCG<br>TGTGCAACGC | ATACTGGTGT<br>TATGACCACA                | TTAGTGCTCC<br>AATCACGAGG | ACTTAAACTC<br>TGAATTTGAG | AGGCACAACC<br>TCCGTGTTGG |
| 23101 | ATCCGCGGCA<br>TAGGCGCCGT | GCTCGGTGAA<br>CGAGCCACTT                | GTTTTCACTC<br>CAAAAGTGAG | CACAGGCTGC<br>GTGTCCGACG | GCACCATCAC<br>CGTGGTAGTG |
| 23151 | CAACGCGTTT<br>GTTGCGCAAA | AGCAGGTCGG<br>TCGTCCAGCC                | GCGCCGATAT<br>CGCGGCTATA | CTTGAAGTCG<br>GAACTTCAGC | CAGTTGGGGC<br>GTCAACCCCG |
| 23201 | GAGGCGGGAC               | CGCGCGCGAG<br>GCGCGCGCTC                | AACGCTATGT               | GTCCCAACGT               | CGTGACCTTG               |
| 23251 | TGATAGTCGC               | CCGGGTGGTG<br>GGCCCACCAC                | GTGCGACCGG               | TCGTGCGAGA               | ACAGCCTCTA               |
| 23301 | GTCTAGGCGC               | TCCAGGTCCT<br>AGGTCCAGGA                | GGCGCAACGA               | GTCCCGCTTG               | CCTCAGTTGA               |
| 23351 | AACCATCGAC               |   | TTCCCGCGCA               | CGGGTCCGAA               | ACTCAACGTG               |
| 23401 |                          | GTGGCATCAA<br>CACCGTAGTT                |                          |                          | GGGCGTTAGG<br>CCCGCAATCC |
| 23451 | TATGTCGCGG               | ACGTATTTTC                              | GGAACTAGAC               | GAATTTTCGG               | ACCTGAGCCT<br>TGGACTCGGA |
| •     | AACGCGGAAG               | TCTCTTCTTG                              | TACGGCGTTC               | TGAACGGCCT               | AAACTGATTG<br>TTTGACTAAC |
|       | CGGCCTGTCC               | GGCGCAGCAC                              | GTGCGTCGTG               | GAACGCAGCC               | TGTTGGAGAT<br>ACAACCTCTA |
|       | GACGTGGTGT               | AAAGCCGGGG                              | TGGCCAAGAA               | GTGCTAGAAC               | GCCTTGCTAG<br>CGGAACGATC |
| 23651 | ACTGCTCCTT<br>TGACGAGGAA | CAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG | TGCCCGTTTT ACGGGCAAAA    | CGCTCGTCAC               | ATCCATTTCA<br>TAGGTAAAGT |

Figure 26 Y

| WO 02/022080 |                          |                              | PCT/US01/28861 |
|--------------|--------------------------|------------------------------|----------------|
| 23701        | ATCACGTGCT<br>TAGTGCACGA | CATAATGCTT<br>GTATTACGAA     |                |
| 23751        | GCCTTCGATC<br>CGGAAGCTAG | GGTGCAGCCA<br>CCACGTCGGT     |                |
| 23801        | CGTGATGCTT<br>GCACTACGAA | TCTGCAAACG<br>AGACGTTTGC     |                |
| 23851        | AATCGCCCCA<br>TTAGCGGGGT | AAAGGTCTTG<br>TTTCCAGAAC     |                |
| 23901        | CAACCCGCGG<br>GTTGGGCGCC | TCAGCCAGGT<br>AGTCGGTCCA     |                |
| 23951        | CTTCCACTTG<br>GAAGGTGAAC | <br>AGTTTGAAGT<br>TCAAACTTCA | <br>           |
| 24001        | ACGTGGTACT<br>TGCACCATGA | CGCGCGCGCA<br>GCGCGCGCGT     | <br>           |
| 24051        | CGCAGACACG<br>GCGTCTGTGC | TCAGCGGGTT<br>AGTCGCCCAA     |                |
| 24101        | CCGCTTCGCT<br>GGCGAAGCGA | TCTTCCTCTT<br>AGAAGGAGAA     | <br>           |
| 24151        | ACTGGGTCGT<br>TGACCCAGCA | <br>CCGCCGCACT<br>GGCGGCGTGA | <br>           |
| 24201        | ATGCTTGATT<br>TACGAACTAA | GGTTGCTGAA<br>CCAACGACTT     | <br>           |
| 24251        |                          | <br>CTGTCCACGA<br>GACAGGTGCT |                |
| 24301        | CGCTCGGGCT<br>GCGAGCCCGA | GCGCTTCTTT<br>CGCGAAGAAA     | ·              |
| 24351        | CAAATCCGCC<br>GTTTAGGCGG | <br>ATGGCCGCGG<br>TACCGGCGCC | <br>           |
| 24401        | GCGCGTCTTG<br>CGCGCAGAAC | TCCTCGTCCT<br>AGGAGCAGGA     |                |
| 24451.       | ATCCGCTTTT<br>TAGGCGAAAA | CCGGGGAGGC<br>GGCCCCTCCG     |                |
| 24501        | CGACACGTCC<br>GCTGTGCAGG | GGGGACGTCG<br>CCCCTGCAGC     | <br>           |
| 24551        | CGGGGGTGGT<br>GCCCCCACCA | TCCTCTTCCC<br>AGGAGAAGGG     | <br>           |
| 24601        | TATAGGCAGA<br>ATATCCGTCT | GGAGTCAGTC<br>CCTCAGTCAG     | <br>           |

Figure 262

| 24701  |                          |            |            |                          | GGAAGTGATT<br>CCTTCACTAA |
|--------|--------------------------|------------|------------|--------------------------|--------------------------|
| 24751  |                          |            |            | GACGACGAGG<br>CTGCTGCTCC |                          |
| 24801  |                          |            |            | CAACGCAGAG<br>GTTGCGTCTC |                          |
| 24851  | AACAAGTCGG<br>TTGTTCAGCC |            |            | GCGACTACCT<br>CGCTGATGGA |                          |
| 24901  |                          |            |            | CAGTGCGCCA<br>GTCACGCGGT |                          |
| 24951  |                          |            |            | CGCCATAGCG<br>GCGGTATCGC |                          |
| 25001  |                          |            |            | GCGTACCCCC<br>CGCATGGGGG |                          |
| 25051  |                          |            |            | CTCAACTTCT<br>GAGTTGAAGA |                          |
| 25101  |                          |            |            | CATCTTTTC<br>GTAGAAAAAG  |                          |
| 25151  | TCTATGGGGA               | TAGGACGGCA | CGGTTGGCGT | GCCGAGCGGA<br>CGGCTCGCCT | GTTCGTCGAC               |
| 25201  |                          |            |            | ATCGCCTCGC<br>TAGCGGAGCG |                          |
| 25251  |                          |            |            | CGAGAAGCGC<br>GCTCTTCGCG |                          |
| 25301  |                          |            |            | GTCACTCTGG<br>CAGTGAGACC |                          |
| 25351  | CTTGAGCTCC               | CACTGTTGCG | CGCGGATCGG | GTACTAAAAC<br>CATGATTTTG | CGTCGTAGCT               |
|        | •                        | AAACGGATGG | GCCGTGAATT | GGATGGGGG                | TTCCAGTACT               |
| 25,451 | GCACAGTCAT<br>CGTGTCAGTA |            |            |                          | CCTGGAGAGG<br>GGACCTCTCC |
| 25501  | GATGCAAATT<br>CTACGTTTAA |            |            |                          | CAGTTGGCGA<br>GTCAACCGCT |
| 25551  | CGAGCAGCTA<br>GCTCGTCGAT |            |            |                          | GACTTGGAGG<br>CTGAACCTCC |

7 igure 26 AA

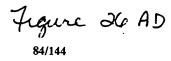
| 25651 |                          | GGTTCTTTGC<br>CCAAGAAACG |              |   | AGCTAGAGGA<br>TCGATCTCCT |
|-------|--------------------------|--------------------------|--------------|---|--------------------------|
| 25701 |                          | TACACCTTTC<br>ATGTGGAAAG |              |   |                          |
| 25751 |                          | GGAGCTCTGC<br>CCTCGAGACG |              | • •                                     |                          |
| 25801 | GAAAACCGCC<br>CTTTTGGCGG | TTGGGCAAAA<br>AACCCGTTTT |              |   | <del>-</del>             |
| 25851 |                          | TACGTCCGCG<br>ATGCAGGCGC | <del>-</del> |   |                          |
| 25901 |                          | CATGGGCGTT<br>GTACCCGCAA |              |   |                          |
| 25951 | AAGGAGCTGC<br>TTCCTCGACG | AGAAACTGCT<br>TCTTTGACGA |              |   |                          |
| 26001 |                          | CGCTCCGTGG<br>GCGAGGCACC | •            |   |                          |
| 26051 |                          | TAAAACCCTG<br>ATTTTGGGAC |              |   |                          |
| 26101 | AGCATGTTGC<br>TCGTACAACG | AGAACTTTAG<br>TCTTGAAATC |              |   |                          |
| 26151 |                          | TGCTGTGCAC<br>ACGACACGTG |              | •                                       |                          |
| 26201 |                          | TCCGCCGCTT<br>AGGCGGCGAA |              |   |                          |
| 26251 |                          | CCTACCACTC<br>GGATGGTGAG |              |   |                          |
| 26301 |                          | TGTCACTGTC<br>ACAGTGACAG |              |   | CACCGCTCCC<br>GTGGCGAGGG |
| 26351 |                          |                          |              | • | CGGTACCTTT<br>GCCATGGAAA |
| 26401 | GAGCTGCAGG<br>CTCGACGTCC | GTCCCTCGCC<br>CAGGGAGCGG |              |   |                          |
| 26451 |                          |                          |              |   | TTTGTACCTG<br>AAACATGGAC |
| 26501 |                          |                          |              |   | ATCCCGCCCG<br>TAGGGCGGGC |

Figure 26 AB

| WO 02/022080 |                          |            |            |            | PCT/US01/28861           |
|--------------|--------------------------|------------|------------|------------|--------------------------|
|              | CCTAATGCGG               | AGCTTACCGC | CTGCGTCATT | ACCCAGGGCC |                          |
| 2000         |                          |            |            |            | TGTAAGAACC               |
| 26601        | CCAATTGCAA               | GCCATCAACA | AAGCCCGCCA | AGAGTTTCTG | CTACGAAAGG               |
|              | GGTTAACGTT               | CGGTAGTTGT | TTCGGGCGGT | TCTCAAAGAC | GATGCTTTCC               |
| 26651        | GACGGGGGGT               | TTACTTGGAC | CCCCAGTCCG | GCGAGGAGCT | CAACCCAATC               |
|              | CTGCCCCCCA               | AATGAACCTG | GGGGTCAGGC | CGCTCCTCGA | GTTGGGTTAG               |
| 26701        | CCCCCGCCGC               |            |            |            |                          |
|              | GGGGGCGCG                | GCGTCGGGAT | AGTCGTCGTC | GGCGCCCGGG | AACGAAGGGT               |
| 26751        | GGATGGCACC               | CAAAAAGAAG | CTGCAGCTGC | CGCCGCCACC | CACGGACGAG               |
|              | CCTACCGTGG               | GTTTTTCTTC | GACGTCGACG | GCGCCGTGG  | GTGCCTGCTC               |
| 26801        | GAGGAATACT               |            |            |            |                          |
|              |                          |            | N          | CAAAACCTGC |                          |
| 26851        | GGACATGATG               |            |            |            |                          |
|              |                          |            |            | GCTCCTTCGA |                          |
| 26901        | AAGAGGTGTC<br>TTCTCCACAG | •          |            |            |                          |
| 26051        | GCGCCCCAGA               |            |            |            |                          |
| 20931        |                          |            |            | TACCGATGTT |                          |
|              |                          |            |            |            |                          |
| 27001        | TCAGGCGCCG               |            |            |            |                          |
|              |                          |            |            | TGGGTTGGCA |                          |
| 27051        | CCACTGGAAC               |            |            | TCGGCGGCG  |                          |
|              | _                        |            |            |            |                          |
| 27101        | GAGCAACAAC               |            |            |            |                          |
|              |                          |            |            | ACCGCGCCCG |                          |
| 27151        | CATAGTTGCT               |            |            | GTTGTAGAGG |                          |
|              |                          |            |            |            |                          |
| 27201        | GCTTTCTTCT               |            |            |            | GTAGGACGTA               |
|              |                          |            |            |            |                          |
| . 27251      |                          |            |            |            | GCGGCAGCAA<br>CGCCGTCGTT |
|              |                          |            |            |            |                          |
| 27301        | CAGCAGCGGC               |            |            |            |                          |
|              |                          |            |            |            | CTGAGACTGT               |
| 27351        | AAGCCCAAGA               |            |            |            |                          |
|              |                          |            |            | •          | CTCGCGACGC               |
| 27401        | TCTGGCGCCC               |            |            |            |                          |
|              |                          |            |            |            | TTGTCCTAAA               |
| 27451        | TTCCCACTCT               |            |            |            | _                        |
|              | AAGGGTGAGA               | CATACGATAT | AAAGTTGTCT | CGTCCCCGGT | TCTTGTTCTC               |

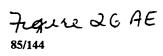
Tiguri 26: AC 83/144

| 27551 |   | GAAGATCAGC<br>CTTCTAGTCG |                             | - |                          |
|-------|---|--------------------------|-----------------------------|---|--------------------------|
| 27601 |   | ATACTGCGCG<br>TATGACGCGC |                             |   |                          |
| 27651 |   | AAGCGCGAAA<br>TTCGCGCTTT |                             |   |                          |
| 27701 |   | GTTGTCAGCG<br>CAACAGTCGC | · · · · · · · · · · · · · · |   |                          |
| 27751 |   | TTACCAGCCA<br>AATGGTCGGT |                             |   |                          |
| 27801 |   | CCCGAATAAA<br>GGGCTTATTT |                             |   |                          |
| 27851 |   | GGAATACGCG<br>CCTTATGCGC |                             |   |                          |
| 27901 |   | CACCACACCT<br>GTGGTGTGGA |                             |   |                          |
| 27951 |   | TGTACCAGGA<br>ACATGGTCCT |                             |   |                          |
| 28001 |   | CAGGCCGAAG<br>GTCCGGCTTC |                             |   |                          |
| 28051 |   | TCGTCACAGG<br>AGCAGTGTCC |                             |   |                          |
| 28101 | • | GAGGGCGAGG<br>CTCCCGCTCC |                             |   |                          |
| 28151 |   |                          |                             |   | CCGCGGCCGGCC             |
| 28201 |   | CACGCCTCGT<br>GTGCGGAGCA |                             |   | GACCTCGTCC<br>CTGGAGCAGG |
| 28251 |   |                          |                             |   | TTGAGGAGTT<br>AACTCCTCAA |
| 28301 |   |                          |                             |   | GGCCACTATC<br>CCGGTGATAG |
| 28351 |   |                          |                             |   | GGCGGACGGC<br>CCGCCTGCCG |
| 28401 |   |                          |                             |   | TGAAACACCT<br>ACTTTGTGGA |



| 02/022080 |                          |            |                          |                          | PCT/US01/28861 |
|-----------|--------------------------|------------|--------------------------|--------------------------|----------------|
| 28451     | CCDCCACDCD               | CCCCCCACA  | እርጥርርጥጥጥርር               | CCGCGACTCC               | CCTCACTTT      |
| 28431     |                          |            |                          | GGCGCTGAGG               |                |
| 20501     | ~~~» ~~~~~~              | אתתיכיריים | CATCATATC                | AGGGCCCGGC               | CCACCCCCTC     |
| 28501     |                          |            |                          | TCCCGGGCCG               |                |
| 28551     | CGGCTTACCG               | CCCAGGGAGA | GCTTGCCCGT               | AGCCTGATTC               | GGGAGTTTAC     |
|           |                          |            |                          | TCGGACTAAG               |                |
| 28601     | CCAGCGCCCC               | CTGCTAGTTG | AGCGGGACAG               | GGGACCCTGT               | GTTCTCACTG     |
|           |                          |            |                          | CCCTGGGACA               |                |
| 28651     | TGATTTGCAA               | СТСТСТВАС  | CCTGGATTAC               | ATCAAGATCT               | TTGTTGCCAT     |
| 20051     |                          |            |                          | TAGTTCTAGA               |                |
| 28701     | CTCTGTGCTG               | AGTATAATAA | ATACAGAAAT               | TAAAATATAC               | TGGGGCTCCT     |
|           |                          |            |                          | ATTTTATATG               |                |
| 28751     | ATCGCCATCC               | TGTAAACGCC | ACCGTCTTCA               | CCCGCCCAAG               | CAAACCAAGG     |
|           |                          |            |                          | GGGCGGGTTC               |                |
| 28801     |                          |            |                          | CCCTCTGTGA               |                |
|           |                          |            |                          | GGGAGACACT               |                |
| 28851     |                          |            |                          | GAACCTCTCC               |                |
|           |                          |            |                          | CTTGGAGAGG               |                |
| 28901     |                          |            |                          | CCTGCCGGGA               |                |
| ٠         |                          |            |                          | GGACGGCCCT               |                |
| 28951     |                          |            |                          | CCTGACCGTA               |                |
| •         |                          |            |                          | GGACTGGCAT               | •              |
| 29001     |                          |            |                          | CCAGAACAGG               |                |
|           |                          |            |                          | GGTCTTGTCC               |                |
| 29051     |                          |            |                          | GCAGCTACTG               |                |
|           |                          |            |                          | CGTCGATGAC               |                |
| 29101     |                          |            |                          | TAATTCAGGT               |                |
|           |                          |            |                          | ATTAAGTCCA               |                |
| 29151     | TCGGGGTTGG<br>AGCCCCAACC | GGTTATTCTC | TGTCTTGTGA<br>ACAGAACACT | TTCTCTTTAT<br>AAGAGAAATA | AGAATATGAT     |
| 29201     | ACGCTTCTCT               | GCCTAAGGCT | CGCCGCCTGC               | TGTGTGCACA               | TTTGCATTTA     |
|           | TGCGAAGAGA               | CGGATTCCGA | GCGGCGGACG               | ACACACGTGT               | AAACGTAAAT     |
| 29251     | TTGTCAGCTT               | TTTAAACGCT | GGGGTCGCCA               | CCCAAGATGA               | TTAGGTACAT     |
| •         | AACAGTCGAA               | AAATTTGCGA | CCCCAGCGGT               | GGGTTCTACT               | AATCCATGTA     |
| 29301     | AATCCTAGGT               | TTACTCACCC | TTGCGTCAGC               | CCACGGTACC               | ACCCAAAAGG     |
|           | TTAGGATCCA               | AATGAGTGGG | AACGCAGTCG               | GGTGCCATGG               | TGGGTTTTCC     |
| 29351     | TGGATTTTAA               | GGAGCCAGCC | TGTAATGTTA               | CATTCGCAGC               | TGAAGCTAAT     |
|           |                          |            |                          |                          | ACTTCGATTA     |

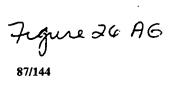
wo



| 29451   |   | AACAAAATTG<br>TTGTTTTAAC | <br> |                          |
|---------|---|--------------------------|------|--------------------------|
| 29501   |   | TACAGAGTAT<br>ATGTCTCATA | <br> |                          |
| 29551   |   | TGTATACTTT<br>ACATATGAAA | <br> |                          |
| 29601   |   | AAACAGTATA<br>TTTGTCATAT | <br> |                          |
| 29651   |   | TTTCTGCTGC<br>AAAGACGACG |      |                          |
| 29701   |   | TACTCTATAT<br>ATGAGATATA |      |                          |
| 29751   | • | ATGCCTTAAT<br>TACGGAATTA | <br> |                          |
| 29801   |   | CTCGCTGCTT<br>GAGCGACGAA |      |                          |
| 29851   |   | GATTTAAACC<br>CTAAATTTGG |      |                          |
| 29901 - | CCTGAACAAT<br>GGACTTGTTA                | TGACTCTATG<br>ACTGAGATAC |      |                          |
| 29951   |   | CTGGATGTCA<br>GACCTACAGT |      |                          |
| 30001   |   | CAGTCCAACT<br>GTCAGGTTGA | <br> |                          |
| 30051   |   | CGCCGCCGCCG              |      |                          |
| 30101   |   | TGCCTTTGTC<br>ACGGAAACAG | <br> | CATGTGGTGG<br>GTACACCACC |
| 30151   |   |                          | <br> | GGCTCATCTG<br>CCGAGTAGAC |
| 30201   |   |                          |      | CCCATCATTG<br>GGGTAGTAAC |
| 30251   | TGCTACACCC<br>ACGATGTGGG                |                          | <br> | ACTGAAACAC<br>TGACTTTGTG |
| 30301   | ATGTTCTTTT<br>TACAAGAAAA                |                          |      | TCCTCGAGTT<br>AGGAGCTCAA |

Figure 26 AF

| 30401  |                          | CACATCGAAG<br>GTGTAGCTTC |                          |                          |                          |
|--------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| 30451  |                          | ATTTGTCACC<br>TAAACAGTGG |                          |                          |                          |
| 30501  |                          | TTATCCAGTG<br>AATAGGTCAC | •                        |                          |                          |
| .30551 |                          | CATCCCCAGT<br>GTAGGGGTCA |                          |                          |                          |
| 30601  |                          | ATTATGAAAT<br>TAATACTTTA |                          |                          |                          |
| 30651  | •                        | GTTTTGTTCC<br>CAAAACAAGG |                          |                          |                          |
| 30701  |                          | CTCGTATATG<br>GAGCATATAC |                          |                          |                          |
| 30751  |                          | GAAGCCTGGT<br>CTTCGGACCA |                          |                          |                          |
| 30801  |                          | CTTAGCCCTA<br>GAATCGGGAT |                          |                          |                          |
| 30851  |                          | ATGCCATGAA<br>TACGGTACTT |                          |                          |                          |
| 30901  |                          | CAAGTTGTTG<br>GTTCAACAAC |                          |                          |                          |
| 30951  |                          | TCCCACCCCC<br>AGGGTGGGGG |                          |                          |                          |
| 31001  |                          | GACACCCTAG<br>CTGTGGGATC |                          |                          |                          |
| 31051  | AGCGCCTGCT<br>TCGCGGACGA | AGAAAGACGC<br>TCTTTCTGCG | AGGGCAGCGG<br>TCCCGTCGCC | CCGAGCAACA<br>GGCTCGTTGT | GCGCATGAAT<br>CGCGTACTTA |
| 31101  |                          |                          |                          |                          | GGGGTATCTT<br>CCCCATAGAA |
| 31151  |                          |                          |                          |                          | ACCACCGGAC<br>TGGTGGCCTG |
| 31201  |                          |                          |                          |                          | GGTGGTCATG<br>CCACCAGTAC |
| 31251  |                          |                          |                          |                          | AAACCGAAGG<br>TTTGGCTTCC |



| 31351 |                          |   | <br>CCTTTAACTA<br>GGAAATTGAT |  |
|-------|--------------------------|---|------------------------------|--|
| 31401 |                          | _ | <br>TAGCAAATTT<br>ATCGTTTAAA |  |
| 31451 |                          |   | AGCTCTGGTA<br>TCGAGACCAT     |  |
| 31501 |                          |   | AATGGAATGT<br>TTACCTTACA     |  |
| 31551 |                          |   | <br>CATGTTGTTG<br>GTACAACAAC |  |
| 31601 |                          |   | CCGTGTATCC<br>GGCACATAGG     |  |
| 31651 |                          |   | <br>ACTCCTCCCT<br>TGAGGAGGGA |  |
| 31701 |                          |   | <br>ACTCTCTTTG<br>TGAGAGAAAC |  |
| 31751 |                          |   | <br>CGCTCAAAAT<br>GCGAGTTTTA |  |
| 31801 |                          |   | TCCCAAAATG<br>AGGGTTTTAC     |  |
| 31851 |                          |   | <br>CATAAACCTG<br>GTATTTGGAC |  |
| 31901 |                          |   | CTGTGGCTGC<br>GACACCGACG     |  |
| 31951 |                          |   | <br>CAATCACAGG<br>GTTAGTGTCC |  |
| 32001 |                          |   | CCAAGGACCC<br>GGTTCCTGGG     |  |
| 32051 | CAGAAGGAAA<br>GTCTTCCTTT |   | GCCCCTCAC<br>CGGGGGAGTG      |  |
| 32101 | AGCAGTACCC<br>TCGTCATGGG |   | CCTCTAACTA<br>GGAGATTGAT     |  |
| 32151 | TAGCTTGGGC<br>ATCGAACCCG |   | <br>TTATACACAA<br>AATATGTGTT |  |
| 32201 | TAGGACTAAA<br>ATCCTGATTT |   | TAACAGACGA<br>ATTGTCTGCT     |  |

Figure 26 AH

| 32301 | AACTAAAGTT<br>TTGATTTCAA                |            | TGGGTTTTGA<br>ACCCAAAACT |            |                          |
|-------|---|------------|--------------------------|------------|--------------------------|
| 32351 | TTAATGTAGC<br>AATTACATCG                |            | AGGATTGATT<br>TCCTAACTAA |            |                          |
| 32401 |   | -          | TGATGCTCAA<br>ACTACGAGTT |            |                          |
| 32451 |   |            | TAAACTCAGC<br>ATTTGAGTCG |            |                          |
| 32501 | ,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,, |            | TTTACAGCTT<br>AAATGTCGAA |            | -                        |
| 32551 |   |            | CAAGGGGTTG<br>GTTCCCCAAC |            |                          |
| 32601 |   |            | GGCTTGAATT<br>CCGAACTTAA |            |                          |
| 32651 |   |            | AAAATTGGCC<br>TTTTAACCGG |            |                          |
| 32701 | TTGTTCCGAT                              | ACCAAGGATT | ACTAGGAACT<br>TGATCCTTGA | CCGGAATCAA | AACTGTCGTG               |
| 32751 |   |            | ACAAAAATAA<br>TGTTTTTATT |            |                          |
| 32801 | GGTGTGGTCG                              | AGGTAGAGGA | AACTGTAGAC<br>TTGACATCTG | ATTTACGTCT | CTTTCTACGA               |
| 32851 |   | ACCAGAATTG | TTTTACACCG               | TCAGTTTATG | AACGATGTCA               |
| 32901 | AAGTCAAAAC                              | CGACAATTTC | GCAGTTTGGC<br>CGTCAAACCG | AGGTTATAGA | CCTTGTCAAG               |
| 32951 | TTTCACGAGT                              | AGAATAATAT |                          | TTTTACCTCA | CGATGATTTG               |
|       | TTAAGGAAGG                              | ACCTGGGTCT | TATAACCTTG               | AAATCTTTAC | GAGATCTTAC<br>CTCTAGAATG |
|       | ACTICCGTGT                              | CGGATATGTT | TGCGACAACC               | TAAATACGGA | AACCTATCAG<br>TTGGATAGTC |
|       | GAATAGGTTT                              | TAGAGTGCCA | TTTTGACGGT               | TTTCATTGTA | TGTCAGTCAA<br>ACAGTCAGTT |
| 33151 |   |            |                          |            | CCATTACACT<br>GGTAATGTGA |

Figure 26 AI 89/144

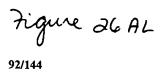
| 33251 |   | GGACTGGTCT<br>CCTGACCAGA |      |                          |
|-------|---|--------------------------|------|--------------------------|
| 33301 |   | ACACTTTTTC<br>TGTGAAAAAG |      |                          |
| 33351 |   | CAACGTGTTT<br>GTTGCACAAA |      |                          |
| 33401 |   | GTAGTATAGC<br>CATCATATCG |      |                          |
| 33451 | • | CAAACTCACA<br>GTTTGAGTGT |      |                          |
| 33501 |   | AGAGTACACA<br>TCTCATGTGT |      |                          |
| 33551 |   | GGGTAACAGA<br>CCCATTGTCT | <br> |                          |
| 33601 |   | GCCAAACGCT<br>CGGTTTGCGA |      |                          |
| 33651 |   | GTTCATGTCG<br>CAAGTACAGC |      |                          |
| 33701 |   | GTTGCTTAAC<br>CAACGAATTG |      |                          |
| 33751 |   | TCATAATCGT<br>AGTATTAGCA |      |                          |
| 33801 |   | AAACTGCTGC<br>TTTGACGACG |      | GGAATACAAC<br>CCTTATGTTG |
| 33851 |   | TCTCCTCAGC<br>AGAGGAGTCG |      | GCATAAGGCG<br>CGTATTCCGC |
| 33901 |   |                          |      | AAATCAGCAC<br>TTTAGTCGTG |
| 33951 |   |                          |      | ACAGTGCAAG<br>TGTCACGTTC |
| 34001 |   |                          |      | CGTGGCCATC<br>GCACCGGTAG |
| 34051 |   |                          |      | AACACGCTGG<br>TTGTGCGACC |
| 34101 |   |                          |      | CTCCCGGTAC<br>GAGGGCCATG |

Figure 26 AJ

| 34201 |            | ACCTGCCCGC               |            |            |            |
|-------|------------|--------------------------|------------|------------|------------|
|       |            | TGGACGGGCG               |            |            |            |
| 34251 |            | GTGGAGAGCC               |            |            |            |
|       | TTGTTACTGT | CACCTCTCGG               | GTCCTGAGCA | TTGGTACCTA | GTAGTACGAG |
| 34301 | ••••       | CAATGTTGGC               |            |            |            |
|       |            | GTTACAACCG               |            |            |            |
| 34351 |            | AGCTCCTCCC               |            |            |            |
|       |            | TCGAGGAGGG               |            |            |            |
| 34401 |            | CAGCGTAAAT               |            |            |            |
|       |            | GTCGCATTTA               |            |            |            |
| 34451 |            | GCATTGTCAA               |            |            |            |
|       | +:         | CGTAACAGTT               |            |            |            |
| 34501 |            | GTAGCGCGGG               |            |            |            |
|       | •          | CATCGCGCCC               |            |            |            |
| 34551 |            | AGTGCGCCGA               |            |            |            |
|       |            | TCACGCGGCT               |            |            |            |
|       | ATGCCAAATG |                          |            |            |            |
|       | TACGGTTTAC |                          |            |            |            |
| 34651 |            | ACAAACAGAT               |            |            |            |
|       |            | TGTTTGTCTA               |            |            |            |
| 34701 |            | AGTTGTAGTA               |            |            |            |
| 24753 |            | TCAACATCAT               |            |            |            |
| 34751 |            | GGTTCTATGT<br>CCAAGATACA |            |            |            |
| 2.001 |            | CGCAGAATAA               |            |            |            |
| 34801 |            | GCGTCTTATT               |            |            |            |
| 24051 |            | ACACGGGAGG               |            | •          |            |
| 34851 |            | TGTGCCCTCC               |            |            |            |
|       | TTTTTTATTC |                          |            |            |            |
| 34901 |            |                          |            |            | TAGATAATTC |
|       | -          |                          |            |            |            |
| 34951 | TGAACGCGCT |                          |            |            | GGTTTCTTGT |
|       |            |                          |            |            |            |
| 35001 | GATAATGGCA |                          |            |            |            |
|       |            |                          |            |            | TCCGTTTGCC |
| 35051 | CCCTCACGTC |                          |            |            |            |
|       | GGGAGTGCAG | GTTCACCTGC               | ATTICCGATT | TGGGAAGTCC | CACTTAGAGG |

Figure 26 AK

| 35151 |   | <br>  | CCGAATATTA<br>GGCTTATAAT |                          |
|-------|---|-------|--------------------------|--------------------------|
| 35201 |   | <br>  | CCTTCAGCCT<br>GGAAGTCGGA |                          |
| 35251 |   |       | AGACCTGTAT<br>TCTGGACATA |                          |
| 35301 |   |       | CGTAGGTCCC<br>GCATCCAGGG |                          |
| 35351 |   | <br>  | GACCAGCGCG<br>CTGGTCGCGC |                          |
| 35401 |   |       | TGATTATGAC<br>ACTAATACTG |                          |
| 35451 |   |       | TAAGCTTGTT<br>ATTCGAACAA |                          |
| 35501 |   | <br>  | ATCAGGCAAA<br>TAGTCCGTTT |                          |
| 35551 |   | <br>  | GCAGATAAAG<br>CGTCTATTTC |                          |
| 35601 |   |       | TTTCTCTCAA<br>AAAGAGAGTT |                          |
| 35651 |   |       | CAAAAAAACA<br>GTTTTTTTGT |                          |
| 35701 | • |       | CCTTATAAGC<br>GGAATATTCG |                          |
| 35751 |   | <br>- | ACTGGTCACC<br>TGACCAGTGG |                          |
| 35801 |   |       | GGAGTCATAA<br>CCTCAGTATT | TGTAAGACTC<br>ACATTCTGAG |
| 35851 |   |       |                          | AAGCGACCGA<br>TTCGCTGGCT |
| 35901 |   |       |                          | CATTACAGCC<br>GTAATGTCGG |
| 35951 |   |       |                          | CATAAACACC<br>GTATTTGTGG |
| 36001 |   |       |                          | TCCAGAACAA<br>AGGTCTTGTT |



| TTTCTTTTGG ATAATTTTTT TGTGGTGAGC TGTGCCGTGG TCGAGTT  | TCA<br>AGT |
|--|------------|
| 36151 GTCACAGTGT AAAAAAAGGGC CAAGTGCAGA GCGAGTATAT ATAGGAC CAGTGTCACA TTTTTTCCCG GTTCACGTCT CGCTCATATA TATCCTG   |            |
| 36201 AAAATGACGT AACGGTTAAA GTCCACAAAA AACACCCAGA AAACCGC. TTTTACTGCA TTGCCAATTT CAGGTGTTTT TTGTGGGTCT TTTGGCG   |            |
| 36251 CGAACCTACG CCCAGAAACG AAAGCCAAAA AACCCACAAC TTCCTCA<br>GCTTGGATGC GGGTCTTTGC TTTCGGTTTT TTGGGTGTTG AAGGAGT |            |
| 36301 CGTCACTTCC GTTTTCCCAC GTTACGTCAC TTCCCATTTT AAGAAAA<br>GCAGTGAAGG CAAAAGGGTG CAATGCAGTG AAGGGTAAAA TTCTTTT |            |
| 36351 CAATTCCCAA CACATACAAG TTACTCCGCC CTAAAACCTA CGTCACC<br>GTTAAGGGTT GTGTATGTTC AATGAGGCGG GATTTTGGAT GCAGTGG |            |
| 36401 CCCGTTCCCA CGCCCCGCGC CACGTCACAA ACTCCACCCC CTCATTA GGGCAAGGGT GCGGGGCGCG GTGCAGTGTT TGAGGTGGGG GAGTAAT    |            |
| PacI   |            |
|  | -          |
| 36451 TATTGGCTTC AATCCAAAAT AAGGTATATT ATTGATGATG TTAATTA ATAACCGAAG TTAGGTTTTA TTCCATATAA TAACTACTAC AATTAAT    |            |
| 36501 ATTCGGATCT GCGACGCGAG GCTGGATGGC CTTCCCCATT ATGATTC<br>TAAGCCTAGA CGCTGCGCTC CGACCTACCG GAAGGGGTAA TACTAAG |            |
|  |            |
| 36551 TCGCTTCCGG CGGCATCGGG ATGCCCGCGT TGCAGGCCAT GCTGTCC  | AGG        |
| AGCGAAGGCC GCCGTAGCCC TACGGGCGCA ACGTCCGGTA CGACAGG  |            |
| 36601 CAGGTAGATG ACGACCATCA GGGACAGCTT CAAGGCCAGC AAAAGGC  | CAG        |
| GTCCATCTAC TGCTGGTAGT CCCTGTCGAA GTTCCGGTCG TTTTCCG  |            |
| 36651 GAACCGTAAA AAGGCCGCGT TGCTGGCGTT TTTCCATAGG CTCCGCC  |            |
| CTTGGCATTT TTCCGGCGCA ACGACCGCAA AAAGGTATCC GAGGCGG  | فافاف      |
| 36701 CTGACGAGCA TCACAAAAAT CGACGCTCAA GTCAGAGGTG GCGAAAC<br>GACTGCTCGT AGTGTTTTTA GCTGCGAGTT CAGTCTCCAC CGCTTTG |            |
| 36751 ACAGGACTAT AAAGATACCA GGCGTTTCCC CCTGGAAGCT CCCTCGT  |            |
| TGTCCTGATA TTTCTATGGT CCGCAAAGGG GGACCTTCGA GGGAGCA  | CGC        |
| 36801 CTCTCCTGTT CCGACCCTGC CGCTTACCGG ATACCTGTCC GCCTTTC  |            |
| GAGAGGACAA GGCTGGGACG GCGAATGGCC TATGGACAGG CGGAAAG  | AGG        |
| 36851 CTTCGGGAAG CGTGGCGCTT TCTCATAGCT CACGCTGTAG GTATCTC<br>GAAGCCCTTC GCACCGCGAA AGAGTATCGA GTGCGACATC CATAGAG |            |
| 36901 TCGGTGTAGG TCGTTCGCTC CAAGCTGGGC TGTGTGCACG AACCCCC AGCCACATCC AGCAAGCGAG GTTCGACCCG ACACACGTGC TTGGGGG    |            |

Figure 26 AM

| 37001 |                                       | <br> | CAGCCACTGG<br>GTCGGTGACC |                          |
|-------|---------------------------------------|------|--------------------------|--------------------------|
| 37051 | · · · · · · ·                         | <br> | GAGTTCTTGA<br>CTCAAGAACT |                          |
| 37101 |                                       | <br> | TGGTATCTGC<br>ACCATAGACG |                          |
| 37151 |                                       | <br> | GCTCTTGATC<br>CGAGAACTAG |                          |
| 37201 |                                       | <br> | TGCAAGCAGC<br>ACGTTCGTCG |                          |
| 37251 | •                                     | <br> | GATCTTTTCT<br>CTAGAAAAGA |                          |
| 37301 | · · · · · · · · · · · · · · · · · · · |      | GGATTTTGGT<br>CCTAAAACCA |                          |
| 37351 |                                       | <br> | AATCAATCTA<br>TTAGTTAGAT |                          |
| 37401 |                                       |      | TTAATCAGTG<br>AATTAGTCAC |                          |
| 37451 |                                       | <br> | AGTTGCCTGA<br>TCAACGGACT |                          |
| 37501 |                                       |      | CATCTGGCCC<br>GTAGACCGGG |                          |
| 37551 |                                       |      | CCAGATTTAT<br>GGTCTAAATA |                          |
| 37601 |                                       | <br> | TGGTCCTGCA<br>ACCAGGACGT |                          |
| 37651 |                                       |      | AAGCTAGAGT<br>TTCGATCTCA |                          |
| 37701 |                                       |      |                          | GCATCGTGGT<br>CGTAGCACCA |
| 37751 |                                       |      |                          | TCCCAACGAT<br>AGGGTTGCTA |
| 37801 | CAAGGCGAGT<br>GTTCCGCTCA              |      | GCAAAAAAGC<br>CGTTTTTTCG |                          |
| 37851 |                                       |      |                          | TGTTATCACT<br>ACAATAGTGA |

Figure 26 AN

| 37951 | GATGCTTTTC<br>CTACGAAAAG |                          | GAGTACTCAA<br>CTCATGAGTT |                          |                          |
|-------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| 38001 |                          |                          | CTCTTGCCCG<br>GAGAACGGGC |                          |                          |
| 38051 |                          |                          | TAAAAGTGCT<br>ATTTTCACGA |                          |                          |
| 38101 | CGGGGCGAAA<br>GCCCCGCTTT |                          | ATCTTACCGC<br>TAGAATGGCG |                          |                          |
| 38151 |                          |                          | CTGATCTTCA<br>GACTAGAAGT |                          |                          |
| 38201 |                          |                          | CAGGAAGGCA<br>GTCCTTCCGT |                          |                          |
| 38251 | TAAGGGCGAC<br>ATTCCCGCTG | ACGGAAATGT<br>TGCCTTTACA | TGAATACTCA<br>ACTTATGAGT | TACTCTTCCT<br>ATGAGAAGGA | TTTTCAATAT<br>AAAAGTTATA |
| 38301 |                          |                          | TTATTGTCTC<br>AATAACAGAG |                          |                          |
| 38351 |                          |                          | AAATAGGGGT<br>TTTATCCCCA |                          |                          |
| 38401 |                          |                          | GAAACCATTA<br>CTTTGGTAAT |                          |                          |
| 38451 |                          |                          | GCCCTTTCGT<br>CGGGAAAGCA |                          | TGGATCCGAA<br>ACCTAGGCTT |
|       |                          | PacI                     |                          |                          |                          |

38501 TTCTTAATTT CTTAATTAA (SEQ ID NO:32) AAGAATTAAA GAATTAATT (SEQ ID NO:33)

Figure 26 AO

| 1   | CATCATCAAT<br>GTAGTAGTTA |      | GAAGCCAATA<br>CTTCGGTTAT |                          |
|-----|--------------------------|------|--------------------------|--------------------------|
| 51  |                          |      | TGGGAACGGG<br>ACCCTTGCCC |                          |
| 101 |                          | <br> | GTGTGGCGGA<br>CACACCGCCT |                          |
| 151 | •                        | <br> | GTGTGCGCCG<br>CACACGCGGC |                          |
| 201 | GAAGTGACAA<br>CTTCACTGTT |      | GATGTTGTAG<br>CTACAACATC |                          |
| 251 |                          |      | GGGAAAACTG<br>CCCTTTTGAC |                          |
| 301 |                          |      | TAGCGCGTAA<br>ATCGCGCATT |                          |
| 351 |                          |      | AGACTCGCCC<br>TCTGAGCGGG |                          |
| 401 | •                        |      | TTGGCGTTTT<br>AACCGCAAAA |                          |
| 451 |                          |      | CATATCATAA<br>GTATAGTATT |                          |
| 501 |                          |      | TGTTGACATT<br>ACAACTGTAA |                          |
| 551 | •                        | <br> | ATTAGTTCAT<br>TAATCAAGTA |                          |
| 601 |                          |      | ATGGCCCGCC<br>TACCGGGCGG |                          |
| 651 |                          | •    | ATGACGTATG<br>TACTGCATAC |                          |
| 701 |                          |      |                          | TATTTACGGT<br>ATAAATGCCA |
| 751 | AAACTGCCCA<br>TTTGACGGGT |      |                          | AAGTACGCCC<br>TTCATGCGGG |
| 801 |                          |      |                          | ATGCCCAGTA<br>TACGGGTCAT |

Figure 27A

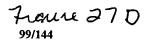
| 901  | • •        | CATGGTGATG<br>GTACCACTAC |            |            |                          |
|------|------------|--------------------------|------------|------------|--------------------------|
| 951  |            | ACTCACGGGG<br>TGAGTGCCCC |            |            |                          |
| 1001 |            | TTTTGGCACC<br>AAAACCGTGG |            |            |                          |
| 1051 |            | CCCATTGACG<br>GGGTAACTGC |            |            |                          |
| 1101 | CAGATATATT | GCAGAGCTCG<br>CGTCTCGAGC | AAATCACTTG | GCAGTCTAGC | GGACCTCTGC               |
| 1151 | GGTAGGTGCG | TGTTTTGACC<br>ACAAAACTGG | AGGTATCTTC | TGTGGCCCTG | GCTAGGTCGG               |
| 1201 |            | CCTTGCCACG               | TAACCTTGCG | CCTAAGGGGC | ACGGTTCTCA               |
| 1251 | CTCTAGACGG | ACCATGGCCG<br>TGGTACCGGC | CGTTCACCAG | GTTCTCCAGG | CACGGGCCGA               |
| 1301 | CCAGGTGGCA | GAGGGAGAGG<br>CTCCCTCTCC | TACTCCTCCC | GGCTCGGGCG | GCGGCTGTCC               |
| 1351 | CACTCCTCCT | CCGAGCCCGC<br>GGCTCGGGCG | GCGTCACCCG | CACCCGCGGC | ACAGGTCCCT               |
| 1401 |            | CACGGCGCCA<br>GTGCCGCGGT |            |            |                          |
| 1451 |            | CTGGCTGGAG<br>GACCGACCTC |            |            |                          |
| 1501 |            | AGGTGCCCCT<br>TCCACGGGGA |            |            |                          |
| 1551 |            | TTCCTGAAGG<br>AAGGACTTCC |            |            |                          |
|      | GGGTCTTCTC | CGTCCTGTAG               | GACCTGGACA | CCCACATGGT | CACCCAGGGC<br>GTGGGTCCCG |
| 1651 | TACTTCCCCG |                          |            |            | TCAGGTTCCC<br>AGTCCAAGGG |
| 1701 |            |                          |            |            | CCCGAGAAGG<br>GGGCTCTTCC |
| 1751 |            |                          |            |            | CCCCATGTCC<br>GGGGTACAGG |

Figure 27B

| 1851 |                          |   | <br>GGAGCTGCAC<br>CCTCGACGTG |                          |
|------|--------------------------|---|------------------------------|--------------------------|
| 1901 |                          |   | <br>CTGTGCCTTC<br>GACACGGAAG |                          |
| 1951 |                          |   | <br>TCCTTGACCC<br>AGGAACTGGG |                          |
| 2001 |                          |   | <br>GGAAATTGCA<br>CCTTTAACGT |                          |
| 2051 |                          |   | <br>GGGTGGGGCA<br>CCCACCCGT  |                          |
| 2101 |                          |   | <br>GCTGGGGATG<br>CGACCCCTAC |                          |
| 2151 |                          |   | <br>TGTGGGCGTG<br>ACACCCGCAC |                          |
| 2201 | -, -                     |   | GTAGTTTTGT<br>CATCAAAACA     | ATCTGTTTTG<br>TAGACAAAAC |
| 2251 |                          |   | <br>CGTTTGATGG<br>GCAAACTACC |                          |
| 2301 |                          |   | <br>TGGGCCGGGG<br>ACCCGGCCCC |                          |
| 2351 |                          |   | CGTCCTGCCC<br>GCAGGACGGG     |                          |
| 2401 |                          | • | <br>CGCCGTTGGA<br>GCGGCAACCT |                          |
| 2451 |                          |   | <br>GCCCGCGGGA<br>CGGGCGCCCT |                          |
| 2501 |                          |   | TGCAGCTTCC<br>ACGTCGAAGG     | •                        |
| 2551 | CCCGCGATGA<br>GGGCGCTACT |   | CACAATTGGA<br>GTGTTAACCT     |                          |
| 2601 | CGGGAACTTA<br>GCCCTTGAAT |   |                              | GCCAGCAGGT<br>CGGTCGTCCA |
| 2651 | TTCTGCCCTG<br>AAGACGGGAC |   | TGCGGTTTAA<br>ACGCCAAATT     |                          |
| 2701 |                          |   |                              | TTGCTGTCTT<br>AACGACAGAA |

Figure 27C

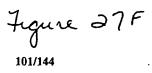
| 2751 | TATTTAGGGG<br>ATAAATCCCC | TTTTGCGCGC<br>AAAACGCGCG | GCGGTAGGCC<br>CGCCATCCGG | CGGGACCAGC<br>GCCCTGGTCG | GGTCTCGGTC<br>CCAGAGCCAG |
|------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| 2801 |                          | CTGTGTATTT<br>GACACATAAA |                          |                          |                          |
| 2851 |                          | CATGGGCATA<br>GTACCCGTAT |                          |                          |                          |
| 2901 |                          | CATGCTGCGG<br>GTACGACGCC |                          |                          |                          |
| 2951 |                          | GCGTGGTGCC<br>CGCACCACGG |                          |                          |                          |
| 3001 |                          | GCCCTTGGTG<br>CGGGAACCAC |                          |                          |                          |
| 3051 |                          | GTGGGGATAT<br>CACCCCTATA |                          |                          |                          |
| 3101 |                          | CCAGCCATAT<br>GGTCGGTATA |                          |                          |                          |
| 3151 | CCAGCACAGT<br>GGTCGTGTCA | GTATCCGGTG<br>CATAGGCCAC | CACTTGGGAA<br>GTGAACCCTT | ATTTGTCATG<br>TAAACAGTAC | TAGCTTAGAA<br>ATCGAATCTT |
| 3201 |                          | GGAAGAACTT<br>CCTTCTTGAA |                          |                          |                          |
| 3251 |                          | TCCATAATGA<br>AGGTATTACT |                          |                          |                          |
| 3301 | CGAAGATATT<br>GCTTCTATAA | TCTGGGATCA<br>AGACCCTAGT | CTAACGTCAT<br>GATTGCAGTA | AGTTGTGTTC<br>TCAACACAAG | CAGGATGAGA<br>GTCCTACTCT |
| 3351 |                          | CCATTTTTAC<br>GGTAAAAATG |                          |                          |                          |
| 3401 | TATAATGGTT<br>ATATTACCAA | CCATCCGGCC<br>GGTAGGCCGG | CAGGGGCGTA<br>GTCCCCGCAT | GTTACCCTCA<br>CAATGGGAGT | CAGATTTGCA<br>GTCTAAACGT |
| 3451 | TTTCCCACGC<br>AAAGGGTGCG | TTTGAGTTCA<br>AAACTCAAGT | GATGGGGGGA<br>CTACCCCCCT | TCATGTCTAC<br>AGTACAGATG | CTGCGGGGCG<br>GACGCCCCGC |
| 3501 | ATGAAGAAAA<br>TACTTCTTTT | CGGTTTCCGG<br>GCCAAAGGCC | GGTAGGGGAG<br>CCATCCCCTC | ATCAGCTGGG<br>TAGTCGACCC | AAGAAAGCAG<br>TTCTTTCGTC |
| 3551 | GTTCCTGAGC<br>CAAGGACTCG | AGCTGCGACT<br>TCGACGCTGA | TACCGCAGCC<br>ATGGCGTCGG | GGTGGGCCCG               | TAAATCACAC<br>ATTTAGTGTG |
| 3601 | CTATTACCGG<br>GATAATGGCC | CTGCAACTGG<br>GACGTTGACC | TAGTTAAGAG<br>ATCAATTCTC | AGCTGCAGCT<br>TCGACGTCGA | GCCGTCATCC<br>CGGCAGTAGG |
| 3651 | CTGAGCAGGG<br>GACTCGTCCC | GGGCCACTTC<br>CCCGGTGAAG | GTTAAGCATG<br>CAATTCGTAC | TCCCTGACTC<br>AGGGACTGAG | GCATGTTTTC<br>CGTACAAAAG |



| 3701 |              | TCCGCCAGAA<br>AGGCGGTCTT |            |            |                          |
|------|--------------|--------------------------|------------|------------|--------------------------|
| 3751 | <del>-</del> | AAAGTTTTTC<br>TTTCAAAAAG |            |            |                          |
| 3801 |              | TTTGACCAAG<br>AAACTGGTTC |            |            |                          |
| 3851 |              | GCATCTCGAT<br>CGTAGAGCTA |            |            |                          |
| 3901 |              | CTGTACGGCA<br>GACATGCCGT |            |            |                          |
| 3951 |              | CCACGGGCGC<br>GGTGCCCGCG |            |            |                          |
| 4001 |              | GCGCTCCGGG<br>CGCGAGGCCC |            |            |                          |
| 4051 | CCAGGACGAC   | GTGCTGAAGC<br>CACGACTTCG | CGACGGCCAG | AAGCGGGACG | CGCAGCCGGT               |
| 4101 | CCATCGTAAA   | GACCATGGTG<br>CTGGTACCAC | AGTATCAGGT | CGGGGAGGCG | CCGCACCGGG               |
| 4151 |              | GCTTGCCCTT<br>CGAACGGGAA |            |            |                          |
| 4201 | TGAAAACTCC   | GCGTAGAGCT<br>CGCATCTCGA | ACCCGCGCTC | TTTATGGCTA | AGGCCCTCA                |
| 4251 | TCCGTAGGCG   | CGGCGTCCGG               | GGCGTCTGCC | AGAGCGTAAG |                          |
| 4301 | CACTCGAGAC   | GCCGTTCGGG<br>CGGCAAGCCC | CAGTTTTTGG | TCCAAAGGGG | GTACGAAAAA               |
| 4351 | CTACGCAAAG   | TTACCTCTGG<br>AATGGAGACC | AAAGGTACTC | GGCCACAGGT | GCGAGCCACT               |
|      | GCTTTTCCGA   | CAGGCACAGG               | GGCATATGTC | TGAACTCTCC |                          |
|      |              | GCGCCAGGAG               | GAGCATATCT | TTGAGCCTGG | TGAGACTCTG               |
|      | TTTCCGAGCG   | CAGGTCCGGT               | CGTGCTTCCT | CCGATTCACC | GAGGGGTAGC<br>CTCCCCATCG |
|      | CCAGCAACAG   | GTGATCCCCC               | AGGTGAGCGA | GGTCCCACAC |                          |
| 4601 |              |                          |            |            | TGTAGGCCAC<br>ACATCCGGTG |



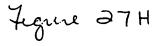
| 4701 |            |            |            | CGAGGGCCAG               |            |
|------|------------|------------|------------|--------------------------|------------|
|      |            |            |            | GCTCCCGGTC               |            |
| 4751 |            |            |            | TCTGCGCTAA               |            |
|      |            |            |            | AGACGCGATT               |            |
| 4801 | TTCCAAAAAC | GAGGAGGATT | TGATATTCAC | CTGGCCCGCG               | GTGATGCCTT |
|      | AAGGTTTTTG | CTCCTCCTAA | ACTATAAGTG | GACCGGGCGC               | CACTACGGAA |
| 4851 |            |            |            | AGACAATCTT               |            |
|      |            |            |            | TCTGTTAGAA               |            |
| 4901 |            |            |            | TTGGACAGCA               |            |
|      |            |            |            | AACCTGTCGT               |            |
| 4951 |            |            |            | GGCGCGCTCC               |            |
|      |            |            |            | CCGCGCGAGG               |            |
| 5001 |            |            |            | ACCGCCATTC               |            |
|      |            |            |            | TGGCGGTAAG               |            |
| 5051 |            |            |            | CGCCAACCGC               |            |
|      |            |            |            | GCGGTTGGCG               |            |
| 5101 |            |            |            | TCCGCGTAGG               |            |
|      |            |            |            | AGGCGCATCC               |            |
| 5151 |            |            |            | AGAATGGCGG               |            |
|      |            |            |            | TCTTACCGCC               |            |
| 5201 |            |            |            | ACGGTAAAGA               |            |
| 5051 |            |            |            | TGCCATTTCT<br>TCCTTGCAAG |            |
| 5251 |            |            |            | AGGAACGTTC               |            |
| 5201 |            |            | -          | CGTATGGGTT               |            |
| 5301 |            |            |            | GCATACCCAA               |            |
|      |            |            |            |                          |            |
| 5351 |            |            |            | GCGTACATGC               |            |
|      |            |            |            | CGCATGTACG               |            |
| 5401 |            |            |            |                          | GGGTAGCATC |
| _    |            |            |            | TTCTATACAT               |            |
| 5451 |            |            |            |                          | GTGCGAGGGA |
|      |            |            |            | GCATATCAAG               |            |
| 5501 | GCGAGGAGGT | CGGGACCGAG | GTTGCTACGG | GCGGGCTGCT               | CTGCTCGGAA |
|      |            |            |            | CGCCCGACGA               |            |
| 5551 |            |            |            |                          | GTTGGACGCT |
|      | CTGATAGACG | GACTTCTACC | GTACACTCAA | CCTACTATAC               | CAACCTGCGA |



| 5651 |   |                  | AGCTCGGCGG<br>TCGAGCCGCC |                          |
|------|---|------------------|--------------------------|--------------------------|
| 5701 |   | <br>             | GATGATGTCA<br>CTACTACAGT |                          |
| 5751 |   | <br>             | GGACAAACTC<br>CCTGTTTGAG |                          |
| 5801 |   | <br>             | GCCTCCGAAC<br>CGGAGGCTTG |                          |
| 5851 |   | <br>             | GGCGCAGCAT<br>CCGCGTCGTA |                          |
| 5901 |   | - · <del>-</del> | GGAGCGAGGT<br>CCTCGCTCCA |                          |
| 5951 | • | <br>             | TACTGGTATT<br>ATGACCATAA |                          |
| 6001 | • | <br>             | AAAGTCCGTG<br>TTTCAGGCAC | -                        |
| 6051 |   | <br>             | CGTTGAAGAG<br>GCAACTTCTC |                          |
| 6101 |   |                  | AAGGGTCCCG<br>TTCCCAGGGC |                          |
| 6151 |   |                  | GATCTCGTCA<br>CTAGAGCAGT |                          |
| 6201 |   | <br>             | AGCGCGGGAT<br>TCGCGCCCTA | = -                      |
| 6251 | • | <br>             | AGCTCTTCAG<br>TCGAGAAGTC |                          |
| 6301 |   |                  | ATGAGGGTTG<br>TACTCCCAAC |                          |
| 6351 |   |                  |                          | GTCGCGAAAG<br>CAGCGCTTTC |
| 6401 |   |                  | TCTGGGGTGA<br>AGACCCCACT | TGCAGTAGAA<br>ACGTCATCTT |
| 6451 |   |                  |                          | GCGGCTAGGT<br>CGCCGATCCA |
| 6501 |   |                  |                          | CATGACCAGC<br>GTACTGGTCG |

Figure 27G

| 6601 |                          | GTGACAAAGA<br>CACTGTTTCT |            |            |                          |
|------|--------------------------|--------------------------|------------|------------|--------------------------|
| 6651 |                          | GATCTCCCGC<br>CTAGAGGGCG | •          |            |                          |
| 6701 |                          | AGTCCCTGCG<br>TCAGGGACGC |            |            |                          |
| 6751 |                          | CAGTACTGGC<br>GTCATGACCG |            |            |                          |
| 6801 | GGTTGACCTG<br>CCAACTGGAC | ACGACCGCGC<br>TGCTGGCGCG |            |            |                          |
| 6851 |                          | GGTTTGGCTG<br>CCAAACCGAC |            |            |                          |
| 6901 | TGGCAGACCG               | TGCTCGAGGG<br>ACGAGCTCCC | CTCAATGCCA | CCTAGCCTOG | TGGTGCGGCG               |
| 6951 |                          | AGTCCAGATG<br>TCAGGTCTAC |            |            |                          |
| 7001 |                          | GATGGGAGCT<br>CTACCCTCGA |            |            |                          |
| 7051 | •                        | AGCTCCTGCA<br>TCGAGGACGT |            |            |                          |
| 7101 |                          | CAGGTGATAC<br>GTCCACTATG |            |            |                          |
| 7151 |                          | GCAAGAGGCC<br>CGTTCTCCGG |            |            |                          |
| 7201 |                          | TGGGCCGCGG<br>ACCCGGCGCC |            |            |                          |
| 7251 |                          | CGAGCCCCCG<br>GCTCGGGGGC |            |            |                          |
| 7301 | GAGGGGGCAG<br>CTCCCCCGTC |                          |            |            | CTGGTGCTGC<br>GACCACGACG |
| 7351 | GCGCGTAGGT<br>CGCGCATCCA |                          |            |            | TCTCCTGAAT<br>AGAGGACTTA |
| 7401 | CTGGCGCCTC               |                          |            |            | AACCTGAAAG<br>TTGGACTTTC |
| 7451 | AGAGTTCGAC<br>TCTCAAGCTG |                          |            |            | CTGGCGCAAA<br>GACCGCGTTT |



| 7551 |                          | TCTTCCTCCT<br>AGAAGGAGGA |            |                          |                          |
|------|--------------------------|--------------------------|------------|--------------------------|--------------------------|
| 7601 |                          | GTCGTTGGAA<br>CAGCAACCTT |            |                          |                          |
| 7651 |                          | CGTTCCAGAC<br>GCAAGGTCTG |            |                          |                          |
| 7701 |                          | ATGACCACCT<br>TACTGGTGGA |            |                          |                          |
| 7751 |                          | GTTTCGCAGG<br>CAAAGCGTCC |            |                          |                          |
| 7801 |                          | CCACGAAGAA<br>GGTGCTTCTT |            |                          |                          |
| 7851 |                          | CCCAAGGCCT<br>GGGTTCCGGA |            |                          |                          |
| 7901 |                          | GAAAAACTGG<br>CTTTTTGACC |            |                          |                          |
| 7951 |                          | GGATGAGCTC<br>CCTACTCGAG |            |                          |                          |
| 8001 |                          | GCCTCTTCTT<br>CGGAGAAGAA |            |                          |                          |
| 8051 |                          | TTCTTCTGGC<br>AAGAAGACCG |            |                          |                          |
| 8101 |                          | CCGGGAGGCG<br>GGCCCTCCGC |            |                          |                          |
| 8151 |                          | ATGGTCTCGG<br>TACCAGAGCC |            |                          |                          |
| 8201 |                          | GCCGCCCGTC<br>CGGCGGGCAG |            |                          |                          |
| 8251 |                          |                          |            |                          | ATTGTTGTGT<br>TAACAACACA |
| 8301 |                          |                          |            |                          | ACCGGATCGG<br>TGGCCTAGCC |
| 8351 |                          |                          |            |                          | AGGTAGGCTG<br>TCCATCCGAC |
| 8401 | AGCACCGTGG<br>TCGTGGCACC | CGGGCGGCAG<br>GCCCGCCGTC | CGGGCGGCGG | TCGGGGTTGT<br>AGCCCCAACA | TTCTGGCGGA<br>AAGACCGCCT |

Figure 27I

| 8501 |                          |                          |                          | CCTGCTGAAT<br>GGACGACTTA |                          |
|------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| 8551 |                          |                          |                          | CGGCGCAGGT<br>GCCGCGTCCA |                          |
| 8601 | GTCTTGCATG<br>CAGAACGTAC | AGCCTTTCTA<br>TCGGAAAGAT | CCGGCACTTC<br>GGCCGTGAAG | TTCTTCTCCT<br>AAGAAGAGGA | TCCTCTTGTC<br>AGGAGAACAG |
| 8651 | CTGCATCTCT<br>GACGTAGAGA | TGCATCTATC<br>ACGTAGATAG | GCTGCGGCGG<br>CGACGCCGCC | CGGCGGAGTT<br>GCCGCCTCAA | TGGCCGTAGG<br>ACCGGCATCC |
| 8701 |                          |                          |                          | CCGAAGCCCC<br>GGCTTCGGGG |                          |
| 8751 |                          |                          |                          | GGCTAATATG<br>CCGATTATAC |                          |
| 8801 |                          |                          |                          | TGTCCACAAA<br>ACAGGTGTTT |                          |
| 8851 | GCGCCCGTGT<br>CGCGGGCACA | TGATGGTGTA<br>ACTACCACAT | AGTGCAGTTG<br>TCACGTCAAC | GCCATAACGG<br>CGGTATTGCC | ACCAGTTAAC<br>TGGTCAATTG |
| 8901 | GGTCTGGTGA<br>CCAGACCACT | CCCGGCTGCG               | AGAGCTCGGT<br>TCTCGAGCCA | GTACCTGAGA<br>CATGGACTCT | CGCGAGTAAG<br>GCGCTCATTC |
| 8951 | CCCTCGAGTC<br>GGGAGCTCAG | AAATACGTAG<br>TTTATGCATC | TCGTTGCAAG<br>AGCAACGTTC | TCCGCACCAG<br>AGGCGTGGTC | GTACTGGTAT<br>CATGACCATA |
| 9001 | CCCACCAAAA<br>GGGTGGTTTT | AGTGCGGCGG<br>TCACGCCGCC | CGGCTGGCGG               | TAGAGGGGCC<br>ATCTCCCCGG | AGCGTAGGGT TCGCATCCCA    |
| 9051 | GGCCGGGGCT<br>CCGGCCCCGA | CCGGGGGCGA               | GATCTTCCAA<br>CTAGAAGGTT | CATAAGGCGA<br>GTATTCCGCT | TGATATCCGT<br>ACTATAGGCA |
| 9101 | AGATGTACCT<br>TCTACATGGA | GGACATCCAG<br>CCTGTAGGTC | GTGATGCCGG<br>CACTACGGCC | CGCCGCCACCA              | GGAGGCGCGC<br>CCTCCGCGCG |
| 9151 | GGAAAGTCGC<br>CCTTTCAGCG | GGACGCGGTT               | CCAGATGTTG<br>GGTCTACAAC | CGCAGCGGCA               | AAAAGTGCTC<br>TTTTCACGAG |
| 9201 | CATGGTCGGG<br>GTACCAGCCC | ACGCTCTGGC<br>TGCGAGACCG | CGGTCAGGCG               | CGCGCAATCG               | TTGACGCTCT<br>AACTGCGAGA |
| 9251 | AGACCGTGCA<br>TCTGGCACGT | AAAGGAGAGC<br>TTTCCTCTCG | CTGTAAGCGG<br>GACATTCGCC | GCACTCTTCC               | CACCAGACCA               |
| 9301 | GGATAAATTC<br>CCTATTTAAG | GCAAGGGTAT<br>GCTTCCCATA | CATGGCGGAC<br>GTACCGCCTG | GACCGGGGTT<br>CTGGCCCCAA | CGAGCCCCGT               |
| 9351 | ATCCGGCCGT<br>TAGGCCGGCA | CCGCCGTGAT<br>GGCGGCACTA | CCATGCGGTT<br>GGTACGCCAA | ACCGCCCGCG               | TGTCGAACCC<br>ACAGCTTGGG |

Figure 27J

| 9451  | CCCCCCCCC                |   | GCCACTGGCC<br>CGGTGACCGG     |                          |
|-------|--------------------------|---|------------------------------|--------------------------|
| 9501  | TAAGCGGTTA<br>ATTCGCCAAT |   | <br>AAGTGGCTCG<br>TTCACCGAGC |                          |
| 9551  |                          |   | <br>CGGGACCCCC<br>GCCCTGGGGG |                          |
| 9601  |                          |   | <br>TTTGCCTCCC<br>AAACGGAGGG |                          |
| 9651  | GACCCCGCTT<br>CTGGGGCGAA |   | <br>GGACGAGCCC<br>CCTGCTCGGG |                          |
| 9701  | TTTCCCAGAT<br>AAAGGGTCTA |   | <br>TGCGCCCCCC<br>ACGCGGGGGG |                          |
| 9751  |                          |   | <br>AGGGCACCCT<br>TCCCGTGGGA |                          |
| 9801  |                          |   | <br>TGACGCGGCA<br>ACTGCGCCGT |                          |
| 9851  | ATTACGAACC<br>TAATGCTTGG |   | <br>ACTACCTGGA<br>TGATGGACCT |                          |
| 9901  |                          |   | TCTCCTGAGC<br>AGAGGACTCG     |                          |
| 9951  |                          | • | GTACGTGCCG<br>CATGCACGGC     |                          |
| 10001 |                          |   | <br>AGGAGATGCG<br>TCCTCTACGC |                          |
| 10051 |                          |   | <br>CTGAATCGCG<br>GACTTAGCGC |                          |
| 10101 |                          |   | <br>AACCGGGATT<br>TTGGCCCTAA |                          |
| 10151 | GCGCACACGT<br>CGCGTGTGCA |   | -                            | GCAGACGGTG<br>CGTCTGCCAC |
| 10201 | AACCAGGAGA<br>TTGGTCCTCT |   |                              | TGCGTACGCT<br>ACGCATGCGA |
| 10251 | TGTGGCGCGC<br>ACACCGCGCG |   | <br>                         | TGGGACTTTG<br>ACCCTGAAAC |
| 10301 | TAAGCGCGCT<br>ATTCGCGCGA |   |                              | GGCGCAGCTG<br>CCGCGTCGAC |

Figure 27K

| 10401 |                          |   | AGGGCCGCTG<br>TCCCGGCGAC |                          |                          |
|-------|--------------------------|---|--------------------------|--------------------------|--------------------------|
| 10451 |                          |   | CAGGAGCGCA<br>GTCCTCGCGT |                          |                          |
| 10501 | GTGGCCGCCA               | TCAACTATTC                              | CATGCTTAGC<br>GTACGAATCG | CTGGGCAAGT               | TTTACGCCCG               |
| 10551 | CAAGATATAC               | CATACCCCTT                              | ACGTTCCCAT<br>TGCAAGGGTA | AGACAAGGAG               | GTAAAGATCG               |
| 10601 |                          |   | GCGCTGAAGG<br>CGCGACTTCC |                          |                          |
| 10651 | CTGGGCGTTT<br>GACCCGCAAA | ATCGCAACGA<br>TAGCGTTGCT                | GCGCATCCAC<br>CGCGTAGGTG | AAGGCCGTGA<br>TTCCGGCACT | GCGTGAGCCG<br>CGCACTCGGC |
| 10701 | GCGGCGCGAG               | CTCAGCGACC                              | GCGAGCTGAT<br>CGCTCGACTA | GCACAGCCTG               | CAAAGGGCCC               |
| 10751 |                          |   | GATAGAGAGG<br>CTATCTCTCC |                          |                          |
| 10801 | GGCGCTGACC<br>CCGCGACTGG | TGCGCTGGGC<br>ACGCGACCCG                | CCCAAGCCGA<br>GGGTTCGGCT | CGCGCCCTGG<br>GCGCGGGACC | AGGCAGCTGG<br>TCCGTCGACC |
| 10851 |                          |   | TGGCACCCGC<br>ACCGTGGGCG |                          |                          |
| 10901 | GCGTGGAGGA<br>CGCACCTCCT | ATATGACGAG<br>TATACTGCTC                | GACGATGAGT<br>CTGCTACTCA | ACGAGCCAGA<br>TGCTCGGTCT | GGACGGCGAG<br>CCTGCCGCTC |
| 10951 |                          |   | GATCAGATGA<br>CTAGTCTACT |                          |                          |
| 11001 | GCGCTGCGGG               | CGGCGCTGCA<br>GCCGCGACGT                | GAGCCAGCCG               | TCCGGCCTTA<br>AGGCCGGAAT | ACTCCACGGA<br>TGAGGTGCCT |
| 11051 | CGACTGGCGC               | CAGGTCATGG<br>GTCCAGTACC                | ACCGCATCAT<br>TGGCGTAGTA | GTCGCTGACT<br>CAGCGACTGA | GCGCGCAATC<br>CGCGCGTTAG |
| 11101 | CTGACGCGTT<br>GACTGCGCAA | CCGGCAGCAG<br>GGCCGTCGTC                | CCGCAGGCCA<br>GGCGTCCGGT | ACCGGCTCTC<br>TGGCCGAGAG | CGCAATTCTG<br>GCGTTAAGAC |
| 11151 | GAAGCGGTGG<br>CTTCGCCACC | TCCCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC | CGCAAACCCC               | ACGCACGAGA<br>TGCGTGCTCT | AGGTGCTGGC<br>TCCACGACCG |
| 11201 | GATCGTAAAC               | CCCTGGCCG                               | AAAACAGGGC               | CATCCGGCCC               | GACGAGGCCG<br>CTGCTCCGGC |
| 11251 | GCCTGGTCTA               | CGACGCGCTG                              | CTTCAGCGCG               | TGGCTCGTTA               | CAACAGCGGC               |

Figure 27L

| 11351  |                          | GAGCGCGCGC<br>CTCGCGCGCGC |       |                          |
|--------|--------------------------|---------------------------|-------|--------------------------|
| 11401  |                          | CTTCCTGAGT<br>GAAGGACTCA  |       |                          |
| 11451  |                          | CCAACTTTGT<br>GGTTGAAACA  | <br>  |                          |
| 11501  |                          | GAGGTGTACC<br>CTCCACATGG  |       |                          |
| 11551  |                          | CCTGCAGACC<br>GGACGTCTGG  | <br>  |                          |
| 11601  |                          | GGGGGGTGCG                | <br>• |                          |
| 13.651 |                          | ACGCCCAACT<br>TGCGGGTTGA  | <br>  |                          |
| 11701  |                          | TGGCAGCGTG<br>ACCGTCGCAC  | <br>  |                          |
| 11751  |                          | GCGAGGCCAT<br>CGCTCCGGTA  | <br>  |                          |
| 11801  |                          | ACAAGTGTCA<br>TGTTCACAGT  |       |                          |
| 11851  |                          | AACCCTAAAC<br>TTGGGATTTG  | <br>  |                          |
| 11901  |                          | ACAGTTTAAA<br>TGTCAAATTT  | <br>  |                          |
| 11951  |                          | GTGAGCCTTA<br>CACTCGGAAT  |       |                          |
| 12001  |                          | CATGACCGCG<br>GTACTGGCGC  |       |                          |
| 12051  | AACCGGCCGT<br>TTGGCCGGCA | TTATCAACCG<br>AATAGTTGGC  | <br>  |                          |
|        | CGTGAACCCC<br>GCACTTGGGG |                           | <br>  |                          |
| 12151  | CGCCCCTGG<br>GCGGGGGACC  | TTTCTACACC<br>AAAGATGTGG  | <br>  |                          |
| 12201  | GGATTCCTCT<br>CCTAAGGAGA |                           | <br>  | CGCAACCGCA<br>GCGTTGGCGT |

Figure 27 M

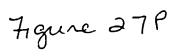
| 12301 |   |      | CCGATCTAGG<br>GGCTAGATCC | CGCTGCGGCC GCGACGCCGG    |
|-------|---|------|--------------------------|--------------------------|
| 12351 |   |      | AGCTTGATAG<br>TCGAACTATC |                          |
| 12401 | • | <br> | GGGCGAGGAG<br>CCCGCTCCTC |                          |
| 12451 |   |      | AAAACCTGCC<br>TTTTGGACGG |                          |
| 12501 |   |      | AAGATGAGTA<br>TTCTACTCAT |                          |
| 12551 |   |      | 922292929<br>0029292929  |                          |
| 12601 |   |      | TGTGGGAGGA<br>ACACCCTCCT |                          |
| 12651 |   |      | GGGAGTGGCA<br>CCCTCACCGT |                          |
| 12701 |   |      | ТТААААААА<br>ААТТТТТТТТ  |                          |
| 12751 |   |      | GCACCGAGCG<br>CGTGGCTCGC |                          |
| 12801 |   | <br> | ATGTATGAGG<br>TACATACTCC |                          |
| 12851 |   |      | GCCAGTGGCG<br>CGGTCACCGC |                          |
| 12901 |   | <br> | CGTTTGTGCC<br>GCAAACACGG |                          |
| 12951 |   |      | CGTTACTCTG<br>GCAATGAGAC |                          |
| 13001 | CCTATTCGAC<br>GGATAAGCTG                |      | GGACAACAAG<br>CCTGTTGTTC |                          |
| 13051 | TGGCATCCCT<br>ACCGTAGGGA                |      |                          | GACCACGGTC<br>CTGGTGCCAG |
| 13101 | ATTCAAAACA<br>TAAGTTTTGT                |      | GCAAGCACAC<br>CGTTCGTGTG |                          |
| 13151 |   |      |                          | ATCCTGCATA<br>TAGGACGTAT |

Figure 27N

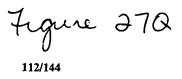
| 13251 |                          | TGTCGCGCTT<br>ACAGCGCGAA |            |              |                          |
|-------|--------------------------|--------------------------|------------|--------------|--------------------------|
| 13301 |                          | GTGGAGTTCA<br>CACCTCAAGT |            |              |                          |
| 13351 |                          | CCTTATGAAC<br>GGAATACTTG |            |              |                          |
| 13401 | · =                      | ACGGGGTTCT<br>TGCCCCAAGA |            |              |                          |
| 13451 |                          | AGACTGGGGT<br>TCTGACCCCA |            |              |                          |
| 13501 |                          | AAACGAAGCC<br>TTTGCTTCGG |            | <del>-</del> |                          |
| 13551 |                          | ACTTCACCCA<br>TGAAGTGGGT |            |              |                          |
| 13601 |                          | CCCTTCCAGG<br>GGGAAGGTCC |            |              |                          |
| 13651 |                          | CATTCCCGCA<br>GTAAGGGCGT |            |              |                          |
| 13701 | · ·                      | ACACCGAACA<br>TGTGGCTTGT |            |              | · ·                      |
| 13751 | *                        | GGCGCGGAAG<br>@CGCGCCTTC |            |              |                          |
| 13801 |                          | GGACATGAAC<br>CCTGTACTTG |            |              |                          |
| 13851 |                          | AGGAGAAGCG<br>TCCTCTTCGC |            | -            |                          |
| 13901 |                          | GCGCAACCCG<br>CGCGTTGGGC |            |              |                          |
|       |                          | СТСТССТС                 | TCGTTCTTTG | CGTCAATGTT   | GGATTATTCG               |
|       |                          | GGAAGTGGGT               | CATGGCGTCG | ACCATGGAAC   | GTATGTTGAT               |
| 14051 | CGGCGACCCT<br>GCCGCTGGGA | CAGACCGGAA<br>GTCTGGCCTT |            |              |                          |
| 14101 | ACGTAACCTG<br>TGCATTGGAC |                          |            |              | AGACATGATG<br>TCTGTACTAC |

Tigure 270

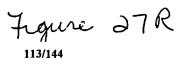
| 14201 |                          | GAGCTGTTGC<br>CTCGACAACG |      |  |
|-------|--------------------------|--------------------------|------|--|
| 14251 |                          | CTCCCAACTC<br>GAGGGTTGAG |      |  |
| 14301 |                          | TTCCCGAGAA<br>AAGGGCTCTT |      |  |
| 14351 |                          | GTCAGTGAAA<br>CAGTCACTTT | <br> |  |
| 14401 |                          | CAACAGCATC<br>GTTGTCGTAG |      |  |
| 14451 |                          | GCACCTGCCC<br>CGTGGACGGG |      |  |
| 14501 |                          | CTATCGAGCC<br>GATAGCTCGG | <br> |  |
| 14551 |                          | CAATAACACA<br>GTTATTGTGT | <br> |  |
| 14601 |                          | CCAAGAAGCG<br>GGTTCTTCGC | <br> |  |
| 14651 |                          | GCGCCCTGGG<br>CGCGGGACCC | <br> |  |
| 14701 |                          | TGACGCCATC<br>ACTGCGGTAG | <br> |  |
| 14751 |                          | CGCCACCAGT<br>GCGGTGGTCA | <br> |  |
| 14801 |                          | GCCCGGCGCT<br>CGGGCCGCGA |      |  |
| 14851 |                          | CCACCGCCGC<br>GGTGGCGGCG | <br> |  |
| 14901 | GCGGCCCTGC<br>CGCCGGGACG | TTAACCGCGC<br>AATTGGCGCG |      |  |
| 14951 | GGCCGCTCGA<br>CCGGCGAGCT | AGGCTGGCCG<br>TCCGACCGGC |      |  |
| 15001 | GGCGACGAGC<br>CCGCTGCTCG | GGCCGCCGCA<br>CCGGCGCGT  |      |  |
| 15051 | GGTCGCAGGG<br>CCAGCGTCCC | GCAACGTGTA<br>CGTTGCACAT |      |  |



| 15151 |   |       |                                       | CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC |                          |
|-------|---|-------|---------------------------------------|---|--------------------------|
| 15201 |   |       |                                       | ATGCTCCAGG<br>TACGAGGTCC                |                          |
| 15251 |   |       |                                       | GCAGGATTAC<br>CGTCCTAATG                |                          |
| 15301 |   |       |                                       | ATGATGATGA<br>TACTACTACT                |                          |
| 15351 | • |       | ••••                                  | CCCAGGCGAC<br>GGGTCCGCTG                |                          |
| 15401 | • |       |                                       | ACCCGGCACC<br>TGGGCCGTGG                |                          |
| 15451 |   |       |                                       | ACAAGCGCGT<br>TGTTCGCGCA                |                          |
| 15501 |   |       |                                       | GCCAACGAGC<br>CGGTTGCTCG                |                          |
| 15551 |   | ••••• |                                       | GCTGGCGTTG<br>CGACCGCAAC                |                          |
| 15601 |   |       |                                       | TAACACTGCA<br>ATTGTGACGT                |                          |
| 15651 |   |       | · · · · · · · · · · · · · · · · · · · | GGCCTAAAGC<br>CCGGATTTCG                |                          |
| 15701 |   |       |                                       | ACCCAAGCGC<br>TGGGTTCGCG                | •                        |
| 15751 |   |       |                                       | CTGGGCTGGA<br>GACCCGACCT                |                          |
| 15801 |   |       |                                       | GGACTGGGCG<br>CCTGACCCGC                |                          |
| 15851 |   |       |                                       | CAGTATTGCC<br>GTCATAACGG                | ACCGCCACAG<br>TGGCGGTGTC |
| 15901 |   |       |                                       |   | GGCGGATGCC<br>CCGCCTACGG |
| 15951 |   |       |                                       | AAGACCTCTA<br>TTCTGGAGAT                |                          |
| 16001 |   |       |                                       |   | CCGCGCCGTT<br>GGCGCGGCAA |



| 16051 |                          |                          |                          | TGCCCGAATA<br>ACGGGCTTAT |                          |
|-------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| 16101 | -                        |                          |                          | GGCTACACCT<br>CCGATGTGGA |                          |
| 16151 |                          |                          |                          | CACTGGAACC<br>GTGACCTTGG |                          |
| 16201 |                          |                          |                          | TTTCCGTGCG<br>AAAGGCACGC |                          |
| 16251 |                          |                          |                          | ACAGCGCGCT<br>TGTCGCGCGA |                          |
| 16301 |                          |                          |                          | TGCAGATATG<br>ACGTCTATAC |                          |
| 16351 |                          |                          |                          | GAGGAAGAAT<br>CTCCTTCTTA |                          |
| 16401 |                          |                          |                          | GGCATGCGTC<br>CCGTACGCAG |                          |
| 16451 |                          |                          |                          | GCGCGGCGGT<br>CGCGCCGCCA |                          |
| 16501 |                          |                          |                          | GCGCCGTGCC<br>CGCGGCACGG |                          |
| 16551 |                          |                          |                          | TTAAAAACAA<br>AATTTTTGTT |                          |
| 16601 |                          |                          |                          | ACGCTCGCTT<br>TGCGAGCGAA |                          |
| 16651 |                          |                          |                          | GCGTCTCTGG<br>CGCAGAGACC |                          |
| 16701 |                          |                          |                          | AGATATCGGC<br>TCTATAGCCG |                          |
| 16751 | TGAGCGGTGG<br>ACTCGCCACC | CGCCTTCAGC<br>GCGGAAGTCG | TGGGGCTCGC<br>ACCCCGAGCG | TGTGGAGCGG<br>ACACCTCGCC | CATTAAAAAT<br>GTAATTTTTA |
| 16801 | TTCGGTTCCA<br>AAGCCAAGGT |                          |                          | AAGGCCTGGA<br>TTCCGGACCT |                          |
| 16851 | AGGCCAGATG<br>TCCGGTCTAC |                          |                          | GCAAAATTTC<br>CGTTTTAAAG |                          |
| 16901 | TGGTAGATGG<br>ACCATCTACC |                          |                          | GGGTGGTGGA<br>CCCACCACCT |                          |
| 16951 | CAGGCAGTGC<br>GTCCGTCACG |                          |                          |                          | GCCCTCCCGT<br>CGGGAGGGCA |



| 17051 |                          |                      | CTCTGGTGAC<br>GAGACCACTG |     |
|-------|--------------------------|----------------------|--------------------------|-----|
| 17161 |                          | <br>                 | CAAGGCCTGC<br>GTTCCGGACG |     |
| 17151 |                          | <br>                 | GGGCCAGCAC<br>CCCGGTCGTG |     |
| 17201 |                          |                      | AGCAGAAACC<br>TCGTCTTTGG |     |
| 17251 |                          | <br>· - <del>-</del> | AGCCGCGCGT<br>TCGGCGCGCA |     |
| 17301 |                          |                      | CGTAGCCAGT<br>GCATCGGTCA |     |
| 17351 |                          | <br>                 | GGGTGCAATC<br>CCCACGTTAG | * * |
| 17401 |                          | <br>                 | ATGTGTGTCA<br>TACACACAGT |     |
| 17451 |                          | <br>                 | GCGCGCGCCC               |     |
| 17501 |                          |                      | TCTTACATGC<br>AGAATGTACG |     |
| 17551 |                          | <br>                 | GCTGGTGCAG<br>CGACCACGTC |     |
| 17601 | · -                      |                      | AGTTTAGAAA<br>TCAAATCTTT |     |
| 17651 |                          |                      | TCCCAGCGTT<br>AGGGTCGCAA |     |
| 17701 |                          |                      | GTACTCGTAC<br>CATGAGCATG |     |
| 17751 | TCACCCTAGC<br>AGTGGGATCG |                      | TGGACATGGC<br>ACCTGTACCG |     |
| 17801 | TTTGACATCC<br>AAACTGTAGG |                      | CCTACTTTTA<br>GGATGAAAAT |     |
| 17851 | TGGCACTGCC<br>ACCGTGACGG |                      | GGGTGCCCCA<br>CCCACGGGGT |     |
| 17901 | AATGGGATGA<br>TTACCCTACT |                      | TAAACCTAGA<br>ATTTGGATCT |     |

Figure 275

| 17951 | GATGACAACG<br>CTACTGTTGC | AAGACGAAGT<br>TTCTGCTTCA | AGACGAGCAA<br>TCTGCTCGTT | GCTGAGCAGC<br>CGACTCGTCG | AAAAAACTCA<br>TTTTTTGAGT |
|-------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| 18001 |                          |                          |                          | AAATATTACA<br>TTTATAATGT |                          |
| 18051 |                          |                          |                          | AATATGCCGA<br>TTATACGGCT |                          |
| 18101 |                          |                          |                          | TGGTACGAAA<br>ACCATGCTTT |                          |
| 18151 |                          |                          |                          | TACCCCAATG<br>ATGGGGTTAC |                          |
| 18201 |                          |                          |                          | ATGGAGGGCA<br>TACCTCCCGT |                          |
| 18251 | CATTTCGTTG               | TTTTACCTTT               | CGATCTTTCA               | CAAGTGGAAA<br>GTTCACCTTT | ACGTTAAAAA               |
| 18301 | CTCAACTACT<br>GAGTTGATGA | GAGGCAGCCG<br>CTCCGTCGGC | CAGGCAATGG<br>GTCCGTTACC | TGATAACTTG<br>ACTATTGAAC | ACTCCTAAAG<br>TGAGGATTTC |
| 18351 |                          |                          |                          | AAACCCCAGA<br>TTTGGGGTCT |                          |
| 18401 |                          |                          |                          | TCACGAGAAC<br>AGTGCTCTTG |                          |
| 18451 |                          |                          |                          | TGCTTTTAGG<br>ACGAAAATCC |                          |
| 18501 | AACCAGATTA               | CATAATGTTG               | TCGTGCCCAT               | ATATGGGTGT<br>TATACCCACA | AGACCGCCCG               |
| 18551 |                          |                          |                          | TTGCAAGACA<br>AACGTTCTGT | GAAACACAGA<br>CTTTGTGTCT |
| 18601 |                          |                          |                          |                          | ACCÀGGTACT<br>TGGTCCATGA |
|       | AAAGATACAC               | ·CTTAGTCCGA              | CAACTGTCGA               | TACTAGGTCT               | TGTTAGAATT<br>ACAATCTTAA |
| 18701 | ATTGAAAATC<br>TAACTTTTAG | ATGGAACTGA<br>TACCTTGACT | AGATGAACTT<br>TCTACTTGAA | CCAAATTACT<br>GGTTTAATGA | GCTTTCCACT<br>CGAAAGGTGA |
| 18751 | GGGAGGTGTG               | ATTAATACAG<br>TAATTATGTC | AGACTCTTAC<br>TCTGAGAATG | CAAGGTAAAA<br>GTTCCATTTT | CCTAAAACAG<br>GGATTTTGTC |
| 18801 | GTCAGGAAAA<br>CAGTCCTTTT | TGGATGGGAA<br>ACCTACCCTT | AAAGATGCTA<br>TTTCTACGAT | CAGAATTTTC<br>GTCTTAAAAG | AGATAAAAAT<br>TCTATTTTA  |
| 18851 | GAAATAAGAG<br>CTTTATTCTC | TTGGAAATAA<br>AACCTTTATT | TTTTGCCATG               | GAAATCAATC               | TAAATGCCAA<br>ATTTACGGTT |

Figure 27T

| 18951 |                          | CAGTCCTTCC<br>GTCAGGAAGG |   |   |  |
|-------|--------------------------|--------------------------|---|---|--|
| 19001 |                          | TGAACAAGCG<br>ACTTGTTCGC |   |   |  |
| 19051 |                          | GGAGCACGCT<br>CCTCGTGCGA |   |   |  |
| 19101 |                          | CCACCGCAAT<br>GGTGGCGTTA |   |   |  |
| 19151 |                          | GCTATGTGCC<br>CGATACACGG |   | • |  |
| 19201 |                          | AACCTCCTTC<br>TTGGAGGAAG |   |   |  |
| 19251 |                          | GGATGTTAAC<br>CCTACAATTG |   |   |  |
| 19301 |                          | ACGGAGCCAG<br>TGCCTCGGTC | • |   |  |
| 19351 |                          | CCCATGGCCC<br>GGGTACCGGG |   |   |  |
| 19401 |                          | CACCAAÇGAC<br>GTGGTTGCTG |   |   |  |
| 19451 |                          | ACCCTATACC<br>TGGGATATGG |   |   |  |
| 19501 |                          | AACTGGGCGG<br>TTGACCCGCC |   |   |  |
| 19551 |                          | AACCCCATCA<br>TTGGGGTAGT |   |   |  |
| 19601 |                          | CTATACCCTA<br>GATATGGGAT |   |   |  |
| 19651 | CTTTAAGAAG<br>GAAATTCTTC | GTGGCCATTA<br>CACCGGTAAT |   |   |  |
| 19701 | ATGACCGCCT<br>TACTGGCGGA | GCTTACCCCC<br>CGAATGGGGG |   |   |  |
| 19751 | GGGGAGGGTT<br>CCCCTCCCAA | ACAACGTTGC<br>TGTTGCAACG |   |   |  |
| 19801 | GGTACAAATG<br>CCATGTTTAC | CTAGCTAACT<br>GATCGATTGA |   |   |  |

Figure 27 4

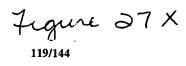
| 19851 |                          |                          |                          | TCTTTAGAAA<br>AGAAATCTTT |                          |
|-------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| 19901 |                          |                          |                          | TACAAGGACT<br>ATGTTCCTGA |                          |
| 19951 |                          |                          |                          | ATTTGTTGGC<br>TAÀACAACCG |                          |
| 20001 |                          |                          |                          | CTAACTTCCC<br>GATTGAAGGG |                          |
| 20051 |                          |                          |                          | CAGAAAAAGT<br>GTCTTTTTCA |                          |
| 20101 | AGCGTGGGAA               | ACCGCGTAGG               | GTAAGAGGTC               | TAACTTTATG<br>ATTGAAATAC | AGGTACCCGC               |
| 20151 | GTGAGTGTCT               | GGACCCGGTT               | TTGGAAGAGA               | ACGCCAACTC<br>TGCGGTTGAG | GCGGGTGCGC               |
| 20201 |                          | GAAAACTCCA               | CCTAGGGTAC               | CTGCTCGGGT               | GGGAAGAAAT               |
| 20251 | ACAAAACAAA               | CTTCAGAAAC               | TGCACCAGGC               | ACACGTGGTC               |                          |
| 20301 | CGCAGTAGCT               | TTGGCACATG               | GACGCGTGCG               | CCTTCTCGGC<br>GGAAGAGCCG | GCCGTTGCGG               |
| 20351 | TGTTGTATTT               | CTTCGTTCGT               | TGTAGTTGTT               | CAGCTGCCGC<br>GTCGACGGCG | GTACCCGAGG               |
| 20401 | TCACTCGTCC               | TTGACTTTCG               | GTAACAGTTT               | GATCTTGGTT<br>CTAGAACCAA | CACCCGGTAT               |
| 20451 | AAAAAACCCG               | TGGATACTGT               | TCGCGAAAGG               | TCCGAAACAA               |                          |
| 20501 | TCGAGCGGAC               | GCGGTATCAG               | TTATGCCGGC               | CAGCGCTCTG               | TGGGGGCGTA<br>ACCCCCGCAT |
|       | GTGACCTACC               | GGAAACGGAC               | CTTGGGCGTG               | AGTTTTTGTA               | GCTACCTCTT<br>CGATGGAGAA |
|       |                          | CCGAAAAGAC               | TGGTCGCTGA               | GTTCGTCCAA               | ATGGTCAAAC               |
|       |                          | TGAGGACGCG               | GCATCGCGGT               | AACGAAGAAG               | GGGGCTGGCG               |
|       | ACATATTGCG               | ACCTTTTCAG               | CTGGGTTTCG               | CATGTCCCCG               | CCAACTCGGC<br>GGTTGAGCCG |
| 20751 | CGCCTGTGGA<br>GCGGACACCT | CTATTCTGCT<br>GATAAGACGA | GCATGTTTCT<br>CGTACAAAGA | CCACGCCTTT               | GCCAACTGGC<br>CGGTTGACCG |

Figure 27 V.

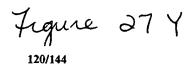
| 20851 |                          | TGCTCAACAG<br>ACGAGTTGTC |                          |                          |                          |
|-------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| 20901 |                          | CTCTACAGCT<br>GAGATGTCGA |                          |                          |                          |
| 20951 | GCCACAGTGC<br>CGGTGTCACG | GCAGATTAGG<br>CGTCTAATCC |                          |                          |                          |
| 21001 |                          | AATGTACTAG<br>TTACATGATC |                          |                          |                          |
| 21051 |                          | CTCGGGTGAT<br>GAGCCCACTA |                          |                          |                          |
| 21101 |                          | AAAGGGGTTC<br>TTTCCCCAAG |                          |                          |                          |
| 21151 | GACACGTTGC<br>CTGTGCAACG | GATACTGGTG<br>CTATGACCAC |                          |                          |                          |
| 21201 |                          | AGCTCGGTGA<br>TCGAGCCACT |                          |                          |                          |
| 21251 |                          | TAGCAGGTCG<br>ATCGTCCAGC |                          |                          |                          |
| 21301 |                          | GCGCGCGCGA<br>CGCGCGCGCT |                          |                          |                          |
| 21351 |                          | GCCGGGTGGT<br>CGGCCCACCA |                          |                          |                          |
| 21401 |                          | GTCCAGGTCC<br>CAGGTCCAGG |                          |                          |                          |
| 21451 |                          |                          |                          |                          | TTGAGTTGCA<br>AACTCAACGT |
| 21501 | GAGCGTGGCA               | TCACCGTAGT               | TTTCCACTGG               | CACGGGCCAG               | TGGGCGTTAG<br>ACCCGCAATC |
| 21551 | GATACAGCGC<br>CTATGTCGCG | CTGCATAAAA<br>GACGTATTTT | GCCTTGATCT<br>CGGAACTAGA | GCTTAAAAGC<br>CGAATTTTCG | CACCTGAGCC<br>GTGGACTCGG |
| 21601 | TTTGCGCCTT               | CAGAGAAGAA<br>GTCTCTTCTT | CATGCCGCAA<br>GTACGGCGTT | GACTTGCCGG<br>CTGAACGGCC | AAAACTGATT<br>TTTTGACTAA |
| 21651 |                          |                          |                          |                          | GTGTTGGAGA<br>CACAACCTCT |
| 21701 | TCTGCACCAC<br>AGACGTGGTG | ATTTCGGCCC<br>TAAAGCCGGG | CACCGGTTCT<br>GTGGCCAAGA | TCACGATCTT<br>AGTGCTAGAA | GGCCTTGCTA<br>CCGGAACGAT |

7, gure 27 W

| 21801 |                          | TCCTTATTTA<br>AGGAATAAAT | · | <br>                         |
|-------|--------------------------|--------------------------|---|------------------------------|
| 21851 |                          | CTCAGCGCAG<br>GAGTCGCGTC |   | <br>                         |
| 21901 |                          | TGTAGGTCAC<br>ACATCCAGTG |   | <br>                         |
| 21951 |                          | ATCATCGTCA<br>TAGTAGCAGT |   | <br>                         |
| 22001 |                          | GTGCTCCTCG<br>CACGAGGAGC |   | <br>                         |
| 22051 |                          | GGTCAGGCAG<br>CCAGTCCGTC |   | <br>                         |
| 22101 |                          | TTGTCCATCA<br>AACAGGTAGT |   |                              |
| 22151 |                          | GATCGGCACA<br>CTAGCCGTGT |   |                              |
| 22201 |                          | TGGGCTCTTC<br>ACCCGAGAAG |   | <br>                         |
| 22251 |                          | TCTTCATTCA<br>AGAAGTAAGT |   | <br>                         |
| 22301 |                          | TAGCACCGGT<br>ATCGTGGCCA |   | <br>                         |
| 22351 |                          | TTTCTTCCTC<br>AAAGAAGGAG |   | <br>                         |
| 22401 |                          | TTGGGAGAAG<br>AACCCTCTTC |   | <br>                         |
|       |                          | CGCCGAGGTC<br>GCGGCTCCAG |   | <br>                         |
| 22501 | AGCGCGTCTT<br>TCGCGCAGAA |                          |   | <br>                         |
| 22551 | CATCCGCTTT<br>GTAGGCGAAA |                          |   | <br>                         |
| 22601 | ACGACACGTC<br>TGCTGTGCAG |                          |   | <br>GCGTCCGCGC<br>CGCAGGCGCG |
| 22651 | TCGGGGGTGG<br>AGCCCCCACC |                          |   |                              |



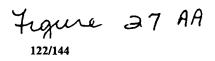
| 22751 |                          | <br>  | CCACCGATGC<br>GGTGGCTACG |                          |
|-------|--------------------------|-------|--------------------------|--------------------------|
| 22801 |                          | <br>  | CTTGAGGAGG<br>GAACTCCTCC |                          |
| 22851 |                          |       | AGACGACGAG<br>TCTGCTGCTC |                          |
| 22901 |                          |       | ACAACGCAGA<br>TGTTGCGTCT |                          |
| 22951 | GAACAAGTCG<br>CTTGTTCAGC |       | GGCGACTACC<br>CCGCTGATGG |                          |
| 23001 |                          | <br>  | CCAGTGCGCC<br>GGTCACGCGG | -                        |
| 23051 |                          | <br>• | TCGCCATAGC<br>AGCGGTATCG |                          |
| 23101 |                          | <br>= | CGCGTACCCC<br>GCGCATGGGG |                          |
| 23151 |                          |       | CCTCAACTTC<br>GGAGTTGAAG |                          |
| 23201 |                          | <br>  | ACATCTTTTT<br>TGTAGAAAAA |                          |
| 23251 |                          | <br>  | AGCCGAGCGG<br>TCGGCTCGCC |                          |
| 23301 |                          | <br>  | TATCGCCTCG<br>ATAGCGGAGC |                          |
| 23351 |                          |       | ACGAGAAGCG<br>TGCTCTTCGC |                          |
| 23401 |                          | <br>  | AGTCACTCTG<br>TCAGTGAGAC |                          |
| 23451 |                          | <br>  | CGTACTAAAA<br>GCATGATTTT | CGCAGCATCG<br>GCGTCGTAGC |
| 23501 |                          |       |                          | CAAGGTCATG<br>GTTCCAGTAC |
| 23551 |                          |       |                          | CCCTGGAGAG<br>GGGACCTCTC |
| 23601 | - <del>-</del>           |       |                          | GCAGTTGGCG<br>CGTCAACCGC |



| 23701 | GAGCGACGCA | AACTAATGAT | GGCCGCAGTG | CTCGTTACCG | TGGAGCTTGA |
|-------|------------|------------|------------|------------|------------|
|       |            | TTGATTACTA |            |            |            |
| 23751 | GTGCATGCAG | CGGTTCTTTG | CTGACCCGGA | GATGCAGCGC | AAGCTAGAGG |
|       |            | GCCAAGAAAC |            |            |            |
| 23801 | AAACATTGCA | CTACACCTTT | CGACAGGGCT | ACGTACGCCA | GGCCTGCAAG |
|       |            | GATGTGGAAA |            |            |            |
| 23851 | ATCTCCAACG | TGGAGCTCTG | CAACCTGGTC | TCCTACCTTG | GAATTTTGCA |
| •     |            | ACCTCGAGAC |            |            | ,          |
| 23901 | CGAAAACCGC | CTTGGGCAAA | ACGTGCTTCA | TTCCACGCTC | AAGGGCGAGG |
|       |            | GAACCCGTTT |            |            |            |
| 23951 | CGCGCCGCGA | CTACGTCCGC | GACTGCGTTT | ACTTATTTCT | ATGCTACACC |
|       |            | GATGCAGGCG |            |            |            |
| 24001 |            | CCATGGGCGT |            |            |            |
|       |            | GGTACCCGCA |            |            |            |
| 24051 | CAAGGAGCTG | CAGAAACTGC | TAAAGCAAAA | CTTGAAGGAC | CTATGGACGG |
|       |            | GTCTTTGACG |            |            |            |
| 24101 | CCTTCAACGA | GCGCTCCGTG | GCCGCGCACC | TGGCGGACAT | CATTTTCCCC |
|       |            | CGCGAGGCAC |            |            |            |
| 24151 |            | TTAAAACCCT |            |            |            |
|       |            | AATTTTGGGA |            |            |            |
| 24201 | AAGCATGTTG | CAGAACTTTA | GGAACTTTAT | CCTAGAGCGC | TCAGGAATCT |
|       |            | GTCTTGAAAT |            |            |            |
| 24251 | TGCCCGCCAC | CTGCTGTGCA | CTTCCTAGCG | ACTTTGTGCC | CATTAAGTAC |
|       |            | GACGACACGT |            |            |            |
| 24301 |            | CTCCGCCGCT |            |            |            |
|       |            |            |            |            | ACGTCGATCG |
| 24351 |            |            |            |            | AGCGGTGACG |
|       |            |            |            |            | TCGCCACTGC |
| 24401 | GTCTACTGGA | GTGTCACTGT | CGCTGCAACC | TATGCACCCC | GCACCGCTCC |
|       | •          |            |            |            | CGTGGCGAGG |
| 24451 | CTGGTTTGCA | ATTCGCAGCT | GCTTAACGAA | AGTCAAATTA | TCGGTACCTT |
|       |            |            |            |            | AGCCATGGAA |
| 24501 | TGAGCTGCAG | GGTCCCTCGC | CTGACGAAAA | GTCCGCGGCT | CCGGGGTTGA |
| •     |            |            |            |            | GGCCCCAACT |
| 24551 | AACTCACTCC | GGGGCTGTGG | ACGTCGGCTT | ACCTTCGCAA | ATTTGTACCT |
|       | TTGAGTGAGG | CCCCGACACC | TGCAGCCGAA | TGGAAGCGTI | TAAACATGGA |

Figure 27Z

| 24601         | GAGGACTACC   | ACGCCCACGA | GATTAGGTTC   | TACGAAGACC           | AATCCCGCCC                              |
|---------------|--------------|------------|--------------|----------------------|---|
|               | · CTCCTGATGG | TGCGGGTGCT | CTAATCCAAG   | ATGCTTCTGG           | TTAGGGCGGG                              |
|               |              |            | •            |                      |   |
| 24651         | GCCTAATGCG   | GAGCTTACCG | CCTGCGTCAT   | TACCCAGGGC           | CACATTCTTG                              |
|               | CGGATTACGC   | CTCGAATGGC | GGACGCAGTA   | ATGGGTCCCG           | GTGTAAGAAC                              |
| •             |              |            |              |                      |   |
| 24701         | GCCAATTGCA   | AGCCATCAAC | AAAGCCCGCC   | AAGAGTTTCT           | GCTACGAAAG                              |
|               | CGGTTAACGT   | TCGGTAGTTG | TTTCGGGCGG   | TTCTCAAAGA           | CGATGCTTTC                              |
|               |              |            |              |                      |   |
| 24751         | GGACGGGGGG   | TTTACTTGGA | CCCCCAGTCC   | GGCGAGGAGC           | TCAACCCAAT                              |
|               | CCTGCCCCCC   | AAATGAACCT | GGGGGTCAGG   | CCGCTCCTCG           | AGTTGGGTTA                              |
|               |              |            |              |                      | •                                       |
| 24801         | CCCCCGCCG    | CCGCAGCCCT | ATCAGCAGCA   | GCCGCGGGCC           | CTTGCTTCCC                              |
|               | GGGGGGCGGC   | GGCGTCGGGA | TAGTCGTCGT   | CGGCGCCCGG           | GAACGAAGGG                              |
|               |              |            |              |                      |   |
| 24851         | AGGATGGCAC   | CCAAAAAGAA | GCTGCAGCTG   | CCGCCGCCAC           | CCACGGACGA                              |
|               |              | GGTTTTTCTT |              |                      |   |
|               |              |            |              |                      |   |
| 24901         | GGAGGAATAC   | TGGGACAGTC | AGGCAGAGGA   | GGTTTTGGAC           | GAGGAGGAGG                              |
|               | CCTCCTTATG   | ACCCTGTCAG | TCCGTCTCCT   | CCAAAACCTG           | CTCCTCCTCC                              |
|               |              |            |              |                      |   |
| 24951         | AGGACATGAT   | GGAAGACTGG | GAGAGCCTAG   | ACGAGGAAGC           | TTCCGAGGTC                              |
|               |              | CCTTCTGACC |              |                      |   |
|               |              |            |              |                      |   |
| 25001         | GAAGAGGTGT   | CAGACGAAAC | ACCGTCACCC   | TCGGTCGCAT           | TCCCCTCGCC                              |
|               |              | GTCTGCTTTG |              |                      |   |
|               |              |            |              |                      |   |
| 25051         | GGCGCCCCAG   | AAATCGGCAA | CCGGTTCCAG   | CATGGCTACA           | ACCTCCCCTC                              |
|               |              | TTTAGCCGTT |              |                      |   |
|               | 2000000000   |            | 00007210010  | omeconio:            | 100/10000                               |
| 25101         | CTCAGGCGCC   | GCCGGCACTG | ררכידדרפרר   | GACCCAACCG           | TAGATGGGAC                              |
|               |              | CGGCCGTGAC |              |                      |   |
|               | 07.0700000   | COOCCOTORC | OGGCANGCGG   | C10001100C           | ATCIACCETO                              |
| 25151         | ACCACTGGAA   | CCAGGGCCGG | таастесаас   | ראפררפררפר           | ССТТАССССА                              |
|               |              | GGTCCCGGCC |              |                      |   |
|               | 100100011    | 001000000  | ATTCAGGITC   | G1CGGCGGCG           | GCANICGGGI                              |
| 25201         | 464664464    | CAGCGCCAAG | CCTACCCCTC   | ATCCCCCCC            | CACAACAACC                              |
| 23201         |              | GTCGCGGTTC |              |                      |   |
|               | 1010011011   | G1C6C6G11C | CGNIGGCGNG   | TACCOCOCCC           | 6161161166                              |
| 25251         | ССАТАСТТСС   | TTGCTTGCAA | GACTGTGGGG   | CCAACATCTC           | רדדרפררפר                               |
|               |              | AACGAACGTT |              |                      |   |
|               | OGIATOMACO   | Ancomed 1  | CIGACACCCC   | CGIIGIAGAG           | amacadaca                               |
| 25301         | CGCTTTCTTC   | ጥሮምልሮሮልጥሮል | רבפרפיזיפפרר | <b>התהכיבות בייו</b> | אראשררשררא                              |
| 23301         |              | AGATGGTAGT |              |                      |   |
|               | GCGMMGMAG    | MONIGOINGI | GCCGCACCGG   | ANGGGGGCN1           | IGINGGACGI                              |
| 25351         | TTACTACCGT   | САТСТСТАСА | GCCCATACTC   | ראַררפפרפפר          | AGCGGCAGCA                              |
| 23331         |              | GTAGAGATGT |              |                      |   |
|               | -WIGHIGGEN   | GINGNONIGI | COGINIGAL    | 31000000000          | .coccorcor                              |
| 25401         | ACAGCAGCGG   | CCACACAGAA | GCAAACGCGA   | CCCCATACCA           | <b>ACACTOTOS</b> AC                     |
| 27401         |              | GGTGTGTCTT |              |                      |   |
|               | 10101000     | GOIGIGICIT | COLLICUSCT   | GGCCTATCGT           | ICIGMGACTG                              |
| 25451         | *******      | 333mcc3c3c | CCCCCCC x CC | **********           | CCXCCCCCCC                              |
| 22421         | AAAGCCCAAG   |            |              |                      |   |
|               | TTTCGGGTTC   | TTTAGGTGTC | GCCGCCGTCG   | TUSTCUTCCT           | CCTCGCGACG                              |
| 25501         | GTCTGGCGCC   | C33CC33CC2 | C#3#CC3 CCC  | OCC NOOME C          | A A A C A C C A C A C A C A C A C A C A |
| <b>7</b> 2201 |              |            |              |                      | <del>-</del>                            |
|               | CAGACCGCGG   | GTTGCTTGGG | CATAGCTGGG   | CGCTCGAATC           | TITGTCCTAA                              |



| O ULIULEU | 00                       |            |            |                          | FC1/030                  |
|-----------|--------------------------|------------|------------|--------------------------|--------------------------|
| 25551     |                          |            |            | AGCAGGGGCC<br>TCGTCCCCGG | AAGAACAAGA<br>TTCTTGTTCT |
| 25601     |                          |            |            | CCTCACCCGC<br>GGAGTGGGCG |                          |
| 25651     |                          |            |            | CGCTGGAAGA<br>GCGACCTTCT |                          |
| 25701     | CTCTTCAGTA<br>GAGAAGTCAT |            |            | AAGGACTAGT<br>TTCCTGATCA |                          |
| 25751     |                          |            |            | TCTCCAGCGG<br>AGAGGTCGCC |                          |
| 25801     |                          |            |            | GCAAGGAAAT<br>CGTTCCTTTA |                          |
| 25851     |                          |            |            | CTTGCGGCTG<br>GAACGCCGAC |                          |
| 25901     |                          |            |            | CGCGGGACCC<br>GCGCCCTGGG |                          |
| 25951     | CCCGGGTCAA               | CGGAATACGC | GCCCACCGAA | ACCGAATTCT<br>TGGCTTAAGA | CCTGGAACAG               |
| 26001     | GCGGCTATTA               | CCACCACACC | TCGTAATAAC | CTTAATCCCC<br>GAATTAGGGG | GTAGTTGGCC               |
| 26051     | CGCTGCCCTG               | GTGTACCAGG | AAAGTCCCGC | TCCCACCACT               | GTGGTACTTC               |
| 26101     | CCAGAGACGC               | CCAGGCCGAA | GTTCAGATGA |                          | GGCGCAGCTT               |
| 26151     |                          |            |            | GATTGAGTCC               | CCGCGTCGAA<br>GTATAACTCA |
| 26201     |                          |            |            | GGGCCCGTCC               | CATATTGAGT               |
|           | GGACTGTTAG               | TCTCCCGCTC | CATAAGTCGA | GTTGCTGCTC               | AGCCACTCGA<br>CGGCGCCGGC |
|           | GGAGCGAACC               | AGAGGCAGGC | CTGCCCTGTA | AAGTCTAGCC               | GCCGCGGCCG<br>AGACCTCGTC |
|           | GCGAGAAGTA               | AGTGCGGAGC | AGTCCGTTAG | GATTGAGACG               | TCTGGAGCAG               |
|           | GAGACTCGGC               | GCGAGACCTC | CGTAACCTTG | AGACGTTAAA               | ATTGAGGAGT<br>TAACTCCTCA |
|           | AACACGGTAG               | CCAGATGAAA | TTGGGGAAGA | GCCCTGGAGG               | CGGCCACTAT               |
| 26451     |                          |            |            |                          | CGGCGGACGG               |

Figure 27 AB

| 26501 |                          | ATGTTAAGTG<br>TACAATTCAC |      |                          |
|-------|--------------------------|--------------------------|------|--------------------------|
| 26551 |                          | TCGCCGCCAC<br>AGCGGCGGTG |      |                          |
| 26601 |                          | AATTGCCCGA<br>TTAACGGGCT | <br> |                          |
| 26651 |                          | GCCCAGGGAG<br>CGGGTCCCTC |      |                          |
| 26701 |                          | CCTGCTAGTT<br>GGACGATCAA | <br> |                          |
| 26751 |                          | ACTGTCCTAA<br>TGACAGGATT |      |                          |
| 26801 |                          | GAGTATAATA<br>CTCATATTAT |      |                          |
| 26851 |                          | CTGTAAACGC<br>GACATTTGCG |      |                          |
| 26901 |                          | CCTGGTACTT<br>GGACCATGAA |      |                          |
| 26951 |                          | AGACGGAGTG<br>TCTGCCTCAC |      |                          |
| 27001 |                          | GAAAAAACAC<br>CTTTTTTGTG |      |                          |
| 27051 |                          | GCCGCTGCAC<br>CGGCGACGTG |      |                          |
| 27101 |                          | AGACCTCAAT<br>TCTGGAGTTA |      |                          |
| 27151 |                          | TTAGGGTATT<br>AATCCCATAA | <br> | GTGGGGTTTA<br>CACCCCAAAT |
| 27201 |                          |                          |      | TTTCTCTAGA<br>AAAGAGATCT |
| 27251 | ATCGGGGTTG<br>TAGCCCCAAC |                          |      | TTCTTATACT<br>AAGAATATGA |
| 27301 | AACGCTTCTC<br>TTGCGAAGAG |                          |      | ATTTGCATTT<br>TAAACGTAAA |
| 27351 | ATTGTCAGCT<br>TAACAGTCGA |                          |      | ATTAGGTACA<br>TAATCCATGT |
| 27401 |                          |                          |      | CACCCAAAAG<br>GTGGGTTTTC |

Figure 27AC

| W | O 02/0220 |                              |   |   |                          | 501/28861 |
|---|-----------|------------------------------|---|---|--------------------------|-----------|
|   | 27451     | AGGAGCCAGC<br>TCCTCGGTCG     |   |   | CTGAAGCTAA<br>GACTTCGATT |           |
|   |           | ACTCTTATAA<br>TGAGAATATT     |   |   |                          |           |
|   | 27551     | <br>AAACAAAATT<br>TTTGTTTTAA |   |   |                          |           |
|   | 27601     | <br>CTACAGAGTA<br>GATGTCTCAT | - |   | -                        |           |
|   | 27651     | ATGTATACTT<br>TACATATGAA     |   |   |                          |           |
|   | 27701     | <br>CAAACAGTAT<br>GTTTGTCATA |   |   |                          |           |
|   | 27751     | <br>CTTTCTGCTG<br>GAAAGACGAC |   |   | TGCTCGCTTT<br>ACGAGCGAAA |           |
|   | 27801     | CTACTCTATA<br>GATGAGATAT     |   |   |                          |           |
|   | 27851     | AATGCCTTAA<br>TTACGGAATT     |   |   |                          |           |
|   | 27901     | <br>ACTCGCTGCT<br>TGAGCGACGA |   |   |                          |           |
|   | 27951     | GGATTTAAAC<br>CCTAAATTTG     |   |   |                          |           |
|   | 28001     | TTGACTCTAT<br>AACTGAGATA     |   |   |                          |           |
|   | 28051     | <br>CCTGGATGTC<br>GGACCTACAG |   |   |                          |           |
|   | 28101     |                              |   |   | GAGATGACCA<br>CTCTACTGGT |           |
|   | 28151     | <br>CCCCCCCCC                |   |   | CACAAATACA<br>GTGTTTATGT |           |
|   | 28201     | <br>CTGCCTTTGT<br>GACGGAAACA |   |   | GCATGTGGTG<br>CGTACACCAC |           |
|   | 28251     | GCGCTTATGT<br>CGCGAATACA     |   |   | TGGCTCATCT<br>ACCGAGTAGA |           |
|   | 28301     | <br>                         |   |   | TCCCATCATT<br>AGGGTAGTAA |           |
|   |           |                              |   | • |                          |           |

Figure 27AD

28351 GTGCTACACC CAAACAATGA TGGAATCCAT AGATTGGACG GACTGAAACA

CACGATGTGG GTTTGTTACT ACCTTAGGTA TCTAACCTGC CTGACTTTGT

| 28451 |   | CTGACCCTTG<br>GACTGGGAAC |              |   |                          |
|-------|---|--------------------------|--------------|---|--------------------------|
| 28501 |   | TCACATCGAA<br>AGTGTAGCTT |              |   |                          |
| 28551 |   | GATTTGTCAC<br>CTAAACAGTG |              |   |                          |
| 28601 |   | TTTATCCAGT<br>AAATAGGTCA |              |   |                          |
| 28651 |   | CCATCCCCAG<br>GGTAGGGGTC |              |   |                          |
| 28701 |   | AATTATGAAA<br>TTAATACTTT |              | • |                          |
| 28751 |   | CGTTTTGTTC<br>GCAAAACAAG | <del>-</del> |   |                          |
| 28801 |   | ACTCGTATAT<br>TGAGCATATA |              |   |                          |
| 28851 |   | CGAAGCCTGG<br>GCTTCGGACC |              |   |                          |
| 28901 |   | TCTTAGCCCT<br>AGAATCGGGA |              |   |                          |
| 28951 |   | GATGCCATGA<br>CTACGGTACT |              |   | -                        |
| 29001 | * | ACAAGTTGTT<br>TGTTCAACAA |              |   |                          |
| 29051 |   | CTCCCACCCC<br>GAGGGTGGGG |              |   |                          |
| 29101 |   | TGACACCCTA<br>ACTGTGGGAT |              |   |                          |
| 29151 |   |                          |              |   | AGCGCATGAA<br>TCGCGTACTT |
| 29201 |   |                          |              |   | AGGGGTATCT<br>TCCCCATAGA |
| 29251 |   |                          |              |   | TACCACCGGA<br>ATGGTGGCCT |
| 29301 |   |                          |              |   | TGGTGGTCAT<br>ACCACCAGTA |

Figure 27 A E

| 29401 | GCTGCATTCA<br>CGACGTAAGT | <br> | AGGATCTCTG<br>TCCTAGAGAC |                          |
|-------|--------------------------|------|--------------------------|--------------------------|
| 29451 |                          |      | CCCTTTAACT<br>GGGAAATTGA |                          |
| 29501 |                          | <br> | TTAGCAAATT<br>AATCGTTTAA | ,                        |
| 29551 |                          | <br> | CAGCTCTGGT<br>GTCGAGACCA |                          |
| 29601 |                          | <br> | AAATGGAATG<br>TTTACCTTAC |                          |
| 29651 |                          |      | TCATGTTGTT<br>AGTACAACAA |                          |
| 29701 |                          |      | CCCGTGTATC<br>GGGCACATAG |                          |
| 29751 |                          |      | TACTCCTCCC<br>ATGAGGAGGG |                          |
| 29801 |                          |      | TACTCTCTTT<br>ATGAGAGAAA |                          |
| 29851 |                          |      | GCGCTCAAAA<br>CGCGAGTTTT |                          |
| 29901 |                          |      | CTCCCAAAAT<br>GAGGGTTTTA |                          |
| 29951 |                          |      | ACATAAACCT<br>TGTATTTGGA |                          |
| 30001 | GCACCCCTCA<br>CGTGGGGAGT |      | ACTGTGGCTG<br>TGACACCGAC |                          |
| 30051 |                          |      | GCAATCACAG<br>CGTTAGTGTC |                          |
| 30101 | CCGTGCACGA<br>GGCACGTGCT |      | CCCAAGGACC<br>GGGTTCCTGG |                          |
| 30151 | TCAGAAGGAA<br>AGTCTTCCTT |      |                          | CCACCACCGA<br>GGTGGTGGCT |
| 30201 | TAGCAGTACC<br>ATCGTCATGG |      | CCCTCTAACT<br>GGGAGATTGA |                          |
| 30251 | GTAGCTTGGG<br>CATCGAACCC |      |                          | AAATGGAAAA<br>TTTACCTTTT |

Lyure 27 AF

| 30351 | TTTGACCGTA<br>AAACTGGCAT | GCAACTGGTC<br>CGTTGACCAG |      |                          |
|-------|--------------------------|--------------------------|------|--------------------------|
| 30401 | AAACTAAAGT               | ·                        |      |                          |
|       |                          | ATGACCTCGG               | <br> |                          |
| 30451 |                          | CAGGAGGACT<br>GTCCTCCTGA |      |                          |
| 30501 |                          | AGTTATCCGT<br>TCAATAGGCA | <br> |                          |
| 30551 |                          | CCCTCTTTTT<br>GGGAGAAAA  |      |                          |
| 30601 |                          |                          | <br> | CCAAAAAGCT<br>GGTTTTTCGA |
| 30651 |                          | CTAAGCACTG<br>GATTCGTGAC | <br> |                          |
| 30701 |                          | TGCAGGAGAT<br>ACGTCCTCTA |      |                          |
| 30751 |                          | CCCTCAAAAC<br>GGGAGTTTTG |      |                          |
| 30801 |                          | ATGGTTCCTA<br>TACCAAGGAT |      |                          |
| 30851 |                          | TACAGTAGGA<br>ATGTCATCCT | <br> |                          |
| 30901 |                          | CTCCATCTCC<br>GAGGTAGAGG | <br> |                          |
| 30951 |                          | TTGGTCTTAA<br>AACCAGAATT | <br> |                          |
| 31001 |                          | GGCTGTTAAA<br>CCGACAATTT |      |                          |
| 31051 | CAAAGTGCTC<br>GTTTCACGAG | ATCTTATTAT<br>TAGAATAATA | <br> | <del> </del>             |
| 31101 | CAATTCCTTC<br>GTTAAGGAAG | CTGGACCCAG<br>GACCTGGGTC |      |                          |
| 31151 | CTGAAGGCAC<br>GACTTCCGTG | AGCCTATACA<br>TCGGATATGT | <br> |                          |
| 31201 | GCTTATCCAA<br>CGAATAGGTT | AATCTCACGG<br>TTAGAGTGCC | <br> |                          |

Figure 27 AG

WO 02/022080 PCT/US01/28861 31251 AGTTTACTTA AACGGAGACA AAACTAAACC TGTAACACTA ACCATTACAC TCAAATGAAT TTGCCTCTGT TTTGATTTGG ACATTGTGAT TGGTAATGTG 31301 TAAACGGTAC ACAGGAAACA GGAGACACAA CTCCAAGTGC ATACTCTATG ATTTGCCATG TGTCCTTTGT CCTCTGTGTT GAGGTTCACG TATGAGATAC 31351 TCATTTTCAT GGGACTGGTC TGGCCACAAC TACATTAATG AAATATTTGC AGTAAAAGTA CCCTGACCAG ACCGGTGTTG ATGTAATTAC TTTATAAACG 31401 CACATCCTCT TACACTTTTT CATACATTGC CCAAGAATAA AGAATCGTTT GTGTAGGAGA ATGTGAAAAA GTATGTAACG GGTTCTTATT TCTTAGCAAA 31451 GTGTTATGTT TCAACGTGTT TATTTTTCAA TTGCAGAAAA TTTCAAGTCA CACAATACAA AGTTGCACAA ATAAAAAGTT AACGTCTTTT AAAGTTCAGT 31501 TTTTCATTC AGTAGTATAG CCCCACCACC ACATAGCTTA TACAGATCAC AAAAAGTAAG TCATCATATC GGGGTGGTGG TGTATCGAAT ATGTCTAGTG 31551 CGTACCTTAA TCAAACTCAC AGAACCCTAG TATTCAACCT GCCACCTCCC GCATGGAATT AGTTTGAGTG TCTTGGGATC ATAAGTTGGA CGGTGGAGGG 31601 TCCCAACACA CAGAGTACAC AGTCCTTTCT CCCCGGCTGG CCTTAAAAAG AGGGTTGTGT GTCTCATGTG TCAGGAAAGA GGGGCCGACC GGAATTTTTC 31651 CATCATATCA TGGGTAACAG ACATATTCTT AGGTGTTATA TTCCACACGG GTAGTATAGT ACCCATTGTC TGTATAAGAA TCCACAATAT AAGGTGTGCC 31701 TTTCCTGTCG AGCCAAACGC TCATCAGTGA TATTAATAAA CTCCCCGGGC AAAGGACAGC TCGGTTTGCG AGTAGTCACT ATAATTATTT GAGGGGCCCG 31751 AGCTCACTTA AGTTCATGTC GCTGTCCAGC TGCTGAGCCA CAGGCTGCTG TCGAGTGAAT TCAAGTACAG CGACAGGTCG ACGACTCGGT GTCCGACGAC 31801 TCCAACTTGC GGTTGCTTAA CGGGCGGCGA AGGAGAAGTC CACGCCTACA AGGTTGAACG CCAACGAATT GCCCGCCGCT TCCTCTTCAG GTGCGGATGT 31851 TGGGGGTAGA GTCATAATCG TGCATCAGGA TAGGGCGGTG GTGCTGCAGC ACCCCCATCT CAGTATTAGC ACGTAGTCCT ATCCCGCCAC CACGACGTCG 31901 AGCGCGCGAA TAAACTGCTG CCGCCGCCGC TCCGTCCTGC AGGAATACAA TCGCGCGCTT ATTTGACGAC GGCGGCGGCG AGGCAGGACG TCCTTATGTT 31951 CATGGCAGTG GTCTCCTCAG CGATGATTCG CACCGCCCGC AGCATAAGGC GTACCGTCAC CAGAGGAGTC GCTACTAAGC GTGGCGGGCG TCGTATTCCG 32001 GCCTTGTCCT CCGGGCACAG CAGCGCACCC TGATCTCACT TAAATCAGCA CGGAACAGGA GGCCCGTGTC GTCGCGTGGG ACTAGAGTGA ATTTAGTCGT 32051 CAGTAACTGC AGCACAGCAC CACAATATTG TTCAAAATCC CACAGTGCAA GTCATTGACG TCGTGTCGTG GTGTTATAAC AAGTTTTAGG GTGTCACGTT 32101 GGCGCTGTAT CCAAAGCTCA TGGCGGGGAC CACAGAACCC ACGTGGCCAT CCGCGACATA GGTTTCGAGT ACCGCCCCTG GTGTCTTGGG TGCACCGGTA 32151 CATACCACAA GCGCAGGTAG ATTAAGTGGC GACCCCTCAT AAACACGCTG

Figure 27 AH

GTATGGTGTT CGCGTCCATC TAATTCACCG CTGGGGAGTA TTTGTGCGAC

| 32251 | CTCTGATTAA<br>GAGACTAATT     |                          |                                    |
|-------|------------------------------|--------------------------|------------------------------------|
| 32301 | AACCTGCCCG<br>TTGGACGGGC     |                          |                                    |
| 32351 | AGTGGAGAGC<br>TCACCTCTCG     |                          | <br>                               |
| 32401 | TCAATGTTGG<br>AGTTACAACC     |                          |                                    |
| 32451 | AAGCTCCTCC<br>TTCGAGGAGG     | · · · <del>-</del> · · - | <br>                               |
| 32501 | <br>TCAGCGTAAA<br>AGTCGCATTT |                          | <br>                               |
| 32551 | TGCATTGTCA<br>ACGTAACAGT     |                          | <br>- · · · · · <del>-</del> · · · |
| 32601 | GGTAGCGCGG<br>CCATCGCGCC     |                          | <br>                               |
| 32651 | <br>GAGTGCGCCG<br>CTCACGCGGC |                          | <br>                               |
| 32701 | GGAACGCCGG<br>CCTTGCGGCC     |                          | <br>                               |
| 32751 | <br>GACAAACAGA<br>CTGTTTGTCT |                          | <br>                               |
| 32801 | <br>TAGTTGTAGT<br>ATCAACATCA |                          | <br>                               |
| 32851 | <br>GGGTTCTATG<br>CCCAAGATAC |                          | <br>                               |
| 32901 | CCGCAGAATA<br>GGCGTCTTAT     |                          | <br>                               |
| 32951 | <br>                         |                          | <br>ACCATGTTTT<br>TGGTACAAA        |
| 33001 | CCAAAAGATT<br>GGTTTTCTAA     |                          | <br>GATCTATTAA<br>CTAGATAATT       |
| 33051 | <br>                         |                          | <br>GCCAAAGAAC<br>CGGTTTCTTG       |
| 33101 | <br>                         |                          | <br>AAGGCAAACG<br>TTCCGTTTGC       |

Figure 27 AI

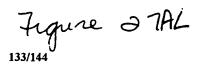
| 33201 | CTCTATAAAC<br>GAGATATTTG |      | GCCCAAATAA<br>CGGGTTTATT |                          |
|-------|--------------------------|------|--------------------------|--------------------------|
| 33251 | GCCACCTTCT<br>CGGTGGAAGA |      | CCCGAATATT<br>GGGCTTATAA |                          |
| 33301 |                          |      | ACCTTCAGCC<br>TGGAAGTCGG |                          |
| 33351 |                          | <br> | CAGACCTGTA<br>GTCTGGACAT |                          |
| 33401 | AAGCGGAACA<br>TTCGCCTTGT |      | CCGTAGGTCC<br>GGCATCCAGG |                          |
| 33451 |                          | <br> | GGACCAGCGC<br>CCTGGTCGCG |                          |
| 33501 |                          |      | CTGATTATGA<br>GACTAATACT |                          |
| 33551 |                          |      | GTAAGCTTGT<br>CATTCGAACA |                          |
| 33601 |                          |      | AATCAGGCAA<br>TTAGTCCGTT |                          |
| 33651 |                          |      | TGCAGATAAA<br>ACGTCTATTT |                          |
| 33701 |                          |      | TTTTCTCTCA<br>AAAAGAGAGT |                          |
| 33751 |                          |      | ACAAAAAAAC<br>TGTTTTTTTG |                          |
| 33801 |                          |      | CCCTTATAAG<br>GGGAATATTC |                          |
| 33851 |                          |      | AACTGGTCAC<br>TTGACCAGTG |                          |
| 33901 | AAGCACCACC<br>TTCGTGGTGG |      |                          | ATGTAAGACT<br>TACATTCTGA |
| 33951 | CGGTAAACAC<br>GCCATTTGTG |      | TCAGTGCTAA<br>AGTCACGATT |                          |
| 34001 | AAATAGCCCG<br>TTTATCGGGC |      | CGTAGAGACA<br>GCATCTCTGT |                          |
| 34051 | CCCCATAGGA<br>GGGGTATCCT |      |                          | ACATAAACAC<br>TGTATTTGTG |

Ligure 27AJ

| 34151 |                          | CTTCCACAGC<br>GAAGGTGTCG |    |      |
|-------|--------------------------|--------------------------|----|------|
| 34201 |                          | CTATTAAAAA<br>GATAATTTTT |    | <br> |
| 34251 |                          | TAAAAAAGGG<br>ATTTTTTCCC |    | <br> |
| 34301 |                          | TAACGGTTAA<br>ATTGCCAATT |    | <br> |
| 34351 |                          | GCCCAGAAAC<br>CGGGTCTTTG |    | <br> |
| 34401 |                          | CGTTTTCCCA<br>GCAAAAGGGT |    | <br> |
| 34451 |                          | ACACATACAA<br>TGTGTATGTT |    | <br> |
| 34501 |                          | ACGCCCCGCG<br>TGCGGGGCGC |    | <br> |
|       |                          |                          |    | PacI |
| 34551 |                          | CAATCCAAAA<br>GTTAGGTTTT |    | <br> |
| 34601 |                          | TGCGACGCGA<br>ACGCTGCGCT | ** |      |
| 34651 |                          | GCGGCATCGG<br>CGCCGTAGCC |    | <br> |
| 34701 |                          | GACGACCATC<br>CTGCTGGTAG |    |      |
| 34751 |                          | AAAGGCCGCG<br>TTTCCGGCGC |    | <br> |
| 34801 | CCTGACGAGC<br>GGACTGCTCG | ATCACAAAAA<br>TAGTGTTTTT |    | <br> |
| 34851 | GACAGGACTA<br>CTGTCCTGAT | TAAAGATACC<br>ATTTCTATGG |    | <br> |
| 34901 | GCTCTCCTGT<br>CGAGAGGACA | TCCGACCCTG<br>AGGCTGGGAC |    | <br> |
| 34951 | CCTTCGGGAA<br>GGAAGCCCTT | GCGTGGCGCT<br>CGCACCGCGA |    |      |
| 35001 | TTCGGTGTAG<br>AAGCCACATC | GTCGTTCGCT<br>CAGCAAGCGA |    |      |

Figure 27 AK

|       | AAGTCGGGCT | GGCGACGCGG               | AATAGGCCAT | TGATAGCAGA | ACTCAGGTTG |
|-------|------------|--------------------------|------------|------------|------------|
| 35101 |            | ACGACTTATC<br>TGCTGAATAG |            |            |            |
| 35151 | TAGCAGAGCG | AGGTATGTAG               | GCGGTGCTAC | AGAGTTCTTG | AAGTGGTGGC |
|       | ATCGTCTCGC | TCCATACATC               | CGCCACGATG | TCTCAAGAAC | TTCACCACCG |
| 35201 |            | CTACACTAGA               |            |            |            |
|       | GATTGATGCC | GATGTGATCT               | TCCTGTCATA | AACCATAGAC | GCGAGACGAC |
| 35251 |            | CCTTCGGAAA               |            |            | •          |
|       | TTCGGTCAAT | GGAAGCCTTT               | TICTCAACCA | TCGAGAACTA | GGCCGTTTGT |
| 35301 |            | GGTAGCGGTG               |            |            |            |
|       | TTGGTGGCGA | CCATCGCCAC               | CAAAAAAACA | AACGTTCGTC | GTCTAATGCG |
| 35351 |            | AGGATCTCAA               |            |            |            |
|       | CGTCTTTTTT | TCCTAGAGTT               | CTTCTAGGAA | ACTAGAAAAG | ATGCCCCAGA |
| 35401 | GACGCTCAGT | GGAACGAAAA               | CTCACGTTAA | GGGATTTTGG | TCATGAGATT |
|       | CTGCGAGTCA | CCTTGCTTTT               | GAGTGCAATT | CCCTAAAACC | AGTACTCTAA |
| 35451 | ATCAAAAAGG | ATCTTCACCT               | AGATCCTTTT | AAATCAATCT | AAAGTATATA |
|       | TAGTTTTTCC | TAGAAGTGGA               | TCTAGGAAAA | TTTAGTTAGA | TTTCATATAT |
| 35501 | TGAGTAAACT | TGGTCTGACA               | GTTACCAATG | CTTAATCAGT | GAGGCACCTA |
|       | ACTCATTTGA | ACCAGACTGT               | CAATGGTTAC | GAATTAGTCA | CTCCGTGGAT |
| 35551 | TCTCAGCGAT | CTGTCTATTT               | CGTTCATCCA | TAGTTGCCTG | ACTCCCCGTC |
|       | AGAGTCGCTA | GACAGATAAA               | GCAAGTAGGT | ATCAACGGAC | TGAGGGGCAG |
| 35601 |            | CTACGATACG               |            |            |            |
|       | CACATCTATT | GATGCTATGC               | CCTCCCGAAT | GGTAGACCGG | GGTCACGACG |
| 35651 |            | CGAGACCCAC               |            |            |            |
|       | TTACTATGGC | GCTCTGGGTG               | CGAGTGGCCG | AGGTCTAAAT | AGTCGTTATT |
| 35701 | ACCAGCCAGC | CGGAAGGGCC               | GAGCGCAGAA | GTGGTCCTGC | AACTTTATCC |
|       | TGGTCGGTCG | GCCTTCCCGG               | CTCGCGTCTT | CACCAGGACG | TTGAAATAGG |
| 35751 |            |                          |            |            | TAAGTAGTTC |
|       | CGGAGGTAGG | TCAGATAATT               | AACAACGGCC | CTTCGATCTC | ATTCATCAAG |
| 35801 | GCCAGTTAAT | AGTTTGCGCA               | ACGTTGTTGC | CATTGCTACA | GGCATCGTGG |
|       | CGGTCAATTA | TCAAACGCGT               | TGCAACAACG | GTAACGATGT | CCGTAGCACC |
| 35851 | TGTCACGCTC | GTCGTTTGGT               | ATGGCTTCAT | TCAGCTCCGG | TTCCCAACGA |
|       |            |                          |            |            | AAGGGTTGCT |
| 35901 | TCAAGGCGAG | TTACATGATC               | CCCCATGTTG | TGCAAAAAAG | CGGTTAGCTC |
|       |            |                          |            |            | GCCAATCGAG |
| 35951 | CTTCGGTCCT | CCGATCGTTG               | TCAGAAGTAA | GTTGGCCGCA | GTGTTATCAC |
|       |            |                          |            |            | CACAATAGTG |
|       |            |                          |            |            |            |



| 36051 | AGATGCTTTT | CTGTGACTGG | TGAGTACTCA | ACCAAGTCAT | TCTGAGAATA  |
|-------|------------|------------|------------|------------|-------------|
|       | TCTACGAAAA | GACACTGACC | ACTCATGAGT | TGGTTCAGTA | AGACTCTTAT  |
| 36101 | GTGTATGCGG | CGACCGAGTT | GCTCTTGCCC | GGCGTCAACA | CGGGATAATA  |
|       | CACATACGCC | GCTGGCTCAA | CGAGAACGGG | CCGCAGTTGT | GCCCTATTAT  |
| 36151 | CCGCGCCACA | TAGCAGAACT | TTAAAAGTGC | TCATCATTGG | AAAACGTTCT  |
|       | GGCGCGGTGT | ATCGTCTTGA | AATTTTCACG | AGTAGTAACC | TTTTGCAAGA  |
| 36201 | TCGGGGCGAA | AACTCTCAAG | GATCTTACCG | CTGTTGAGAT | CCAGTTCGAT  |
|       | AGCCCCGCTT | TTGAGAGTTC | CTAGAATGGC | GACAACTCTA | GGTCAAGCTA  |
| 36251 | GTAACCCACT | CGTGCACCCA | ACTGATCTTC | AGCATCTTTT | ACTTTCACCA  |
|       | CATTGGGTGA | GCACGTGGGT | TGACTAGAAG | TCGTAGAAAA | TGAAAGTGGT  |
| 36301 | GCGTTTCTGG | GTGAGCAAAA | ACAGGAAGGC | AAAATGCCGC | AAAAAAGGGA. |
|       | CGCAAAGACC | CACTCGTTTT | TGTCCTTCCG | TTTTACGGCG | TTTTTTCCCT  |
| 36351 |            | CACGGAAATG |            |            |             |
|       | TATTCCCGCT | GTGCCTTTAC | AACTTATGAG | TATGAGAAGG | AAAAAGTTAT  |
| 36401 |            | ATTTATCAGG |            |            |             |
|       | AATAACTTCG | TAAATAGTCC | CAATAACAGA | GTACTCGCCT | ATGTATAAAC  |
| 36451 | AATGTATTTA | GAAAAATAAA | CAAATAGGGG | TTCCGCGCAC | ATTTCCCCGA  |
|       | TTACATAAAT | CTTTTTATTT | GTTTATCCCC | AAGGCGCGTG | TAAAGGGGCT  |
| 36501 | AAAGTGCCAC | CTGACGTCTA | AGAAACCATT | ATTATCATGA | CATTAACCTA  |
|       | TTTCACGGTG | GACTGCAGAT | TCTTTGGTAA | TAATAGTACT | GTAATTGGAT  |
| 36551 | TAAAAATAGG | CGTATCACGA | GGCCCTTTCG | TCTTCAAGAA | TTGGATCCGA  |
|       | ATTTTTATCC | GCATAGTGCT | CCGGGAAAGC | AGAAGTTCTT | AACCTAGGCT  |
|       |            | PacI       |            |            |             |

36601 ATTCTTAATT TCTTAATTAA (SEQ ID NO:34) TAAGAATTAA AGAATTAATT (SEQ ID NO:35)

Ligure 27 AM

| VIRUS (P5)                                      | PLASMID                                   | VIRUS (P                            | 21)                |
|---|---|-------------------------------------|--------------------|
| MRKAd5gag(E3*)<br>MRKAd5gag(E3*)<br>1 Kb*ladder | pAd5MRKgagSPA(E3*)<br>pAd5MRKmCMVgag(E3*) | MRKAd5gag(E3·)<br>MRKAd5gagSPA(E3·) | MRKAd5mCMVgag(E3*) |
|   |   |                                     |                    |

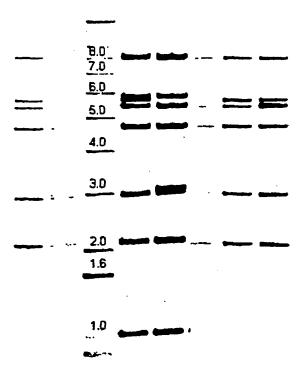


FIGURE 28

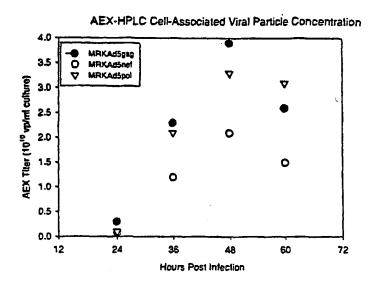


FIGURE 29A

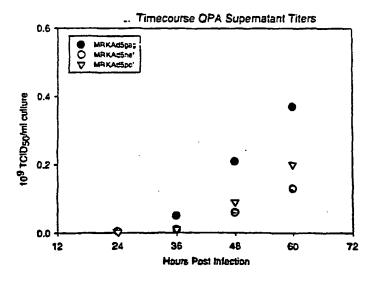


FIGURE 29B

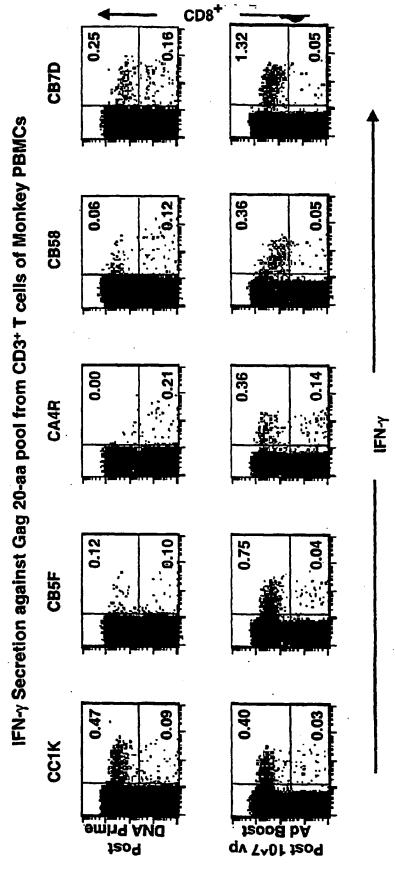
| atg<br>Met<br>1   | gat<br>Asp        | gca<br>Ala        | atg<br>Met        | aag<br>Lys<br>5   | aga<br>Arg        | Gly<br>999        | ctc<br>Leu        | tgc<br>Cys        | tgt<br>Cys<br>10    | gtg<br>Val        | ctg<br>Leu        | ctg<br>Leu               | ctg<br>Leu        | tgt<br>Cys<br>15  | Gly               | 48  |
|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|---------------------|-------------------|-------------------|--------------------------|-------------------|-------------------|-------------------|-----|
| gca<br>Ala        | gtc<br>Val        | ttc<br>Phe        | gtt<br>Val<br>20  | tcg<br>Ser        | ccc<br>Pro        | agc<br>Ser        | gag<br>Glu        | atc<br>Ile<br>25  | tcc<br>Ser          | att<br>Ile        | gtg<br>Val        | tgg<br>Trp               | gcc<br>Ala<br>30  | tcc<br>Ser        | agg<br>Arg        | 96  |
| gag<br>Glu        | ctg<br>Leu        | gag<br>Glu<br>35  | agg<br>Arg        | ttt<br>Phe        | gct<br>Ala        | gtg<br>Val        | aac<br>Asn<br>40  | cct<br>Pro        | ggc                 | ctg<br>Leu        | ctg<br>Leu        | gag<br>Glu<br><b>4</b> 5 | acc<br>Thr        | tct<br>Ser        | gag<br>Glu        | 144 |
| ggg<br>Gly        | tgc<br>Cys<br>50  | agg<br>Arg        | cag<br>Gln        | atc<br>Ile        | ctg<br>Leu        | ggc<br>Gly<br>55  | cag<br>Gln        | ctc<br>Leu        | cag<br>Gln          | ccc<br>Pro        | tcc<br>Ser<br>60  | ctg<br>Leu               | caa<br>Gln        | aca<br>Thr        | GJA<br>BBC        | 192 |
| tct<br>Ser<br>65  | gag<br>Glu        | gag<br>Glu        | ctg<br>Leu        | agg               | tcc<br>Ser<br>70  | ctg<br>Leu        | tac<br>Tyr        | aac<br>Asn        | aca<br>Thr          | gtg<br>Val<br>75  | gct<br>Ala        | acc<br>Thr               | ctg<br>Leu        | tac<br>Tyr        | tgt<br>Cys<br>80  | 240 |
| gtg<br>Val        | cac<br>His        | cag<br>Gln        | aag<br>Lys        | att<br>Ile<br>85  | gat<br>Asp        | gtg<br>Val        | aag<br>Lys        | gac<br>Asp        | acc<br>Thr<br>90    | aag<br>Lys        | gag<br>Glu        | gcc<br>Ala               | ctg<br>Leu        | gag<br>Glu<br>95  | aag<br>Lys        | 288 |
| att<br>Ile        | gag<br>Glu        | gag<br>Glu        | gag<br>Glu<br>100 | cag<br>Gln        | aac<br>Asn        | aag<br>Lys        | tcc<br>Ser        | aag<br>Lys<br>105 | aag<br>Lys          | aag<br>Lys        | gcc<br>Ala        | cag<br>Gln               | cag<br>Gln<br>110 | gct<br>Ala        | gct               | 336 |
| gct<br>Ala        | Gly               | aca<br>Thr<br>115 | G]Å<br>ââc        | aac<br>Asn        | tcc<br>Ser        | agc<br>Ser        | cag<br>Gln<br>120 | gtg<br>Val        | tcc<br>Ser          | cag<br>Gln        | aac<br>Asn        | tac<br>Tyr<br>125        | ccc<br>Pro        | att<br>Ile        | gtg<br>Val        | 384 |
| cag<br>Gln        | aac<br>Asn<br>130 | ctc<br>Leu        | cag<br>Gln        | Gly               | cag<br>Gln        | atg<br>Met<br>135 | gtg<br>Val        | cac<br>His        | cag<br>Gln          | gcc<br>Ala        | atc<br>Ile<br>140 | tcc<br>Ser               | Pro               | cgg               | acc<br>Thr        | 432 |
| ctg<br>Leu<br>145 | Asn               | gcc<br>Ala        | tgg<br>Trp        | gtg<br>Val        | aag<br>Lys<br>150 | Val               | gtg<br>Val        | gag<br>Glu        | gag<br>Glu          | aag<br>Lys<br>155 | Ala               | t t c<br>Phe             | tcc<br>Ser        | cct<br>Pro        | gag<br>Glu<br>160 | 480 |
| gtg<br>Val        | atc<br>Ile        | CCC<br>Pro        | atg<br>Met        | ttc<br>Phe<br>165 | Ser               | Ala               | ctg<br>Leu        | tct<br>Ser        | gag<br>Glu<br>170   | Gly               | gcc<br>Ala        | acc                      | Pro               | cag<br>Gln<br>175 | Asp               | 528 |
| ctg<br>Lev        | aac<br>Asn        | acc<br>Thr        | atg<br>Met<br>180 | Leu               | aac<br>Asn        | aca<br>Thr        | gtg<br>Val        | ggg<br>Gly<br>185 | Gly                 | cat<br>His        | cag<br>Gln        | gct                      | gcc<br>Ala<br>190 | Met               | cag<br>Gln        | 576 |
| atg<br>Met        | ctg<br>Leu        | aag<br>Lys<br>195 | Glu               | acc               | atc<br>Ile        | aat<br>Asn        | gag<br>Glu<br>200 | Glu               | gct<br>Ala          | gct               | gag<br>Glu        | tgg<br>Trp<br>205        | Asp               | agg<br>Arg        | ctg<br>Leu        | 624 |
| cat<br>His        | cct<br>Pro<br>210 | Val               | cac<br>His        | gct               | ggc               | Pro<br>215        | Ile               | gcc<br>Ala        | Pro                 | Gly               | Glr<br>Glr<br>220 | Met                      | agg<br>Arg        | gaç<br>Glu        | Pro               | 672 |
| agg<br>Arg<br>225 | Gly               | tct<br>Ser        | gac<br>Asp        | att<br>Ile        | gct<br>Ala<br>230 | G1y               | acc<br>Thr        | acc<br>Thr        | tcc<br>Ser          | Thr<br>235        | Let               | caç<br>Gli               | gaç<br>Glu        | caç<br>1 Glr      | att<br>lle<br>240 | 720 |
| GJ7<br>BBC        | tgg<br>Trp        | atg<br>Met        | acc<br>Thr        | Asr<br>245        | ASE               | Pro               | ecc<br>Pro        | ato<br>Ile        | e cct<br>Pro<br>250 | Val               | . G17             | g gaa<br>/ Glu           | a ato             | tac<br>Ty:<br>25  | e aag<br>Lys      | 768 |

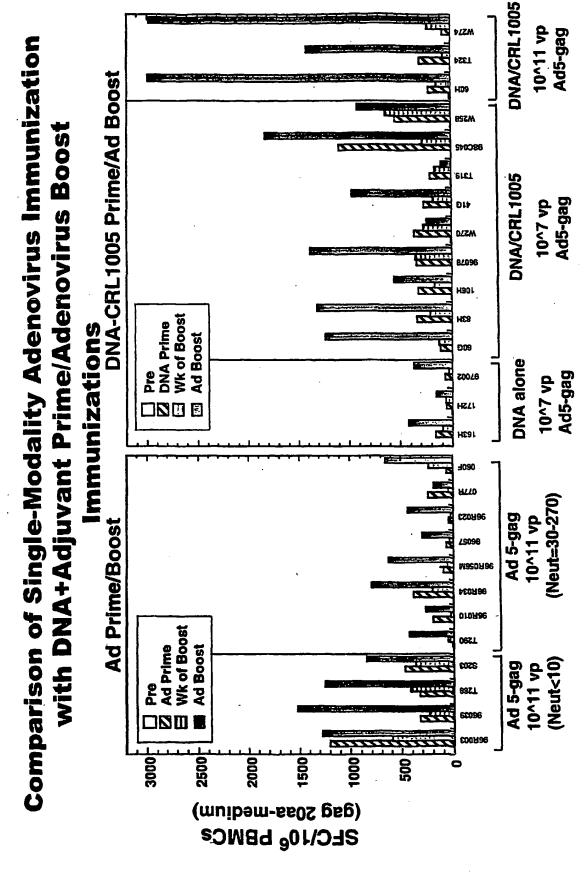
Figure 30'A"

| agg<br>Arg        | tgg<br>Trp        | atc<br>Ile        | atc<br>Ile<br>260 | ctg<br>Leu        | ggc               | ctg<br>Leu         | aac<br>Asn         | aag<br>Lys<br>265  | att<br>Ile        | gtg<br>Val        | agg<br>Arg        | atg<br>Met        | tac<br>Tyr<br>270 | tcc<br>Ser        | ccc<br>Pro        | 816  |
|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|--------------------|--------------------|--------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|------|
| acc<br>Thr        | tcc<br>Ser        | atc<br>Ile<br>275 | ctg<br>Leu        | gac<br>Asp        | atc<br>Ile        | agg<br><b>Ar</b> g | cag<br>Gln<br>280  | ggc<br>Gly         | ccc<br>Pro        | aag<br>Lys        | gag<br>Glu        | ccc<br>Pro<br>285 | ttc<br>Phe        | agg<br>Arg        | gac<br>Asp        | 864  |
| tat<br>Tyr        | gtg<br>Val<br>290 | gac<br>Asp        | agg<br>Arg        | ttc<br>Phe        | tac<br>Tyr        | aag<br>Lys<br>295  | acc                | ctg<br>Leu         | agg<br>Arg        | gct<br>Ala        | gag<br>Glu<br>300 | cag<br>Gln        | gcc<br>Ala        | tcc<br>Ser        | cag<br>Gln        | 912  |
| gag<br>Glu<br>305 | gtg<br>Val        | aag<br>Lys        | aac<br>Asn        | tgg<br>Trp        | atg<br>Met<br>310 | aca<br>Thr         | Gj <i>n</i><br>gag | acc<br>Thr         | ctg<br>Leu        | ctg<br>Leu<br>315 | gtg<br>Val        | cag<br>Gln        | aat<br>Asn        | gcc<br>Ala        | aac<br>Asn<br>320 | 960  |
| cct<br>Pro        | gac<br>Asp        | tgc<br>Cys        | aag<br>Lys        | acc<br>Thr<br>325 | atc<br>Ile        | ctg<br>Leu         | aag<br>Lys         | gcc<br>Ala         | ctg<br>Leu<br>330 | GJÅ<br>ââc        | cct<br>Pro        | gct<br>Ala        | gcc<br>Ala        | acc<br>Thr<br>335 | ctg<br>Leu        | 1008 |
| gag<br>Glu        | gag<br>Glu        | atg<br>Met        | atg<br>Met<br>340 | aca<br>Thr        | gcc<br>Ala        | cys<br>Cys         | cag<br>Gln         | 999<br>Gly<br>345  | gtg<br>Val        | eja<br>aaa        | Gly<br>Gly        | ect<br>Pro        | ggt<br>Gly<br>350 | cac<br>His        | aag<br>Lys        | 1056 |
| gcc<br>Ala        | agg<br>Arg        | gtg<br>Val<br>355 | ctg<br>Leu        | gct<br>Ala        | gag<br>Glu        | gcc<br>Ala         | atg<br>Met<br>360  | tcc<br>Ser         | cag<br>Gln        | gtg<br>Val        | acc<br>Thr        | aac<br>Asn<br>365 | tcc<br>Ser        | gcc<br>Ala        | acc<br>Thr        | 1104 |
| atc<br>Ile        | atg<br>Met<br>370 | atg<br>Met        | cag<br>Gln        | agg<br>Arg        | ggc<br>Gly        | aac<br>Asn<br>375  | ttc<br>Phe         | agg<br><b>A</b> rg | aac<br>Asn        | cag<br>Gln        | agg<br>Arg<br>380 | aag<br>Lys        | aca<br>Thr        | gtg<br>Val        | aag<br>Lys        | 1152 |
| tgc<br>Cys<br>385 | ttc<br>Phe        | aac<br>Asn        | tgt<br>Cys        | ggc               | aag<br>Lys<br>390 | gtg<br>Val         | ggc<br>Gly         | cac<br>His         | att<br>Ile        | gcc<br>Ala<br>395 | aag<br>Lys        | aac<br>Asn        | tgt<br>Cys        | ágg<br>Arg        | gcc<br>Ala<br>400 | 1200 |
| Pro               | agg<br>Arg        | aag<br>Lys        | aag<br>Lys        | ggc<br>Gly<br>405 | tgc<br>Cys        | tgg<br>Trp         | aag<br>Lys         | tgt<br>Cys         | ggc<br>Gly<br>410 | aag<br>Lys        | gag<br>Glu        | ggc<br>Gly        | cac<br>His        | cag<br>Gln<br>415 | atg<br>Met        | 1248 |
| aag<br>Lys        | gac<br>Asp        | tgc<br>Cys        | aat<br>Asn<br>420 | gag<br>Glu        | agg<br>Arg        | cag<br>Gln         | gcc<br>Ala         | aac<br>Asn<br>425  | ttc<br>Phe        | ctg<br>Leu        | GJY<br>ggc        | aaa<br>Lys        | atc<br>Ile<br>430 | tgg<br>Trp        | ccc<br>Pro        | 1296 |
| tcc<br>Ser        | cac<br>His        | aag<br>Lys<br>435 | Gly<br>ggc        | agg<br>Arg        | cct<br>Pro        | Gly<br>ggc         | aac<br>Asn<br>440  | ttc<br>Phe         | ctc<br>Leu        | cag<br>Gln        | tcc<br>Ser        | agg<br>Arg<br>445 | cct<br>Pro        | gag<br>Glu        | ccc<br>Pro        | 1344 |
| aca<br>Thr        | gcc<br>Ala<br>450 | cct<br>Pro        | ccc<br>Pro        | gag<br>Glu        | gag<br>Glu        | tcc<br>Ser<br>455  | ttc<br>Phe         | agg<br>Arg         | ttť<br>Phe        | eja<br>Bab        | gag<br>Glu<br>460 | Glu               | aag<br>Lys        | acc<br>Thr        | acc<br>Thr        | 1392 |
| ccc<br>Pro<br>465 | agc<br>Ser        | cag<br>Gln        | aag<br>Lys        | cag<br>Gln        | gag<br>Glu<br>470 | ccc<br>Pro         | att<br>Ile         | gac<br>Asp         | aag<br>Lys        | gag<br>Glu<br>475 | ctg<br>Leu        | tac<br>Tyr        | ccc<br>Pro        | ctg<br>Leu        | gcc<br>Ala<br>480 | 1440 |
| tcc<br>Ser        | ctg<br>Leu        | agg<br>Arg        | tcc<br>Ser        | ctg<br>Leu<br>485 | ttt<br>Phe        | ggc<br>Gly         | aac<br>Asn         | gac<br>Asp         | ccc<br>Pro<br>490 | tcc<br>Ser        | tcc<br>Ser        | cag<br>Gln        | taa<br>*          |                   | NO:36)            | 1482 |

Figure 30 B

Figure 31





### FIGURE 33A

| NINCCCTYCCTEN | GGGCTTCTGT | こしからかくかららか | GGTGAGCTGG | ACAAGTGGGA | GAAGATCAGG |
|---------------|------------|------------|------------|------------|------------|
|               | GTGGCAAGAA |            |            |            |            |
|               | TTGCTGTGAA |            |            |            |            |
|               | TCCAGCCCTC |            |            |            |            |
|               | CCCTGTACTG |            |            |            |            |
|               | TTGAGGAGGA |            |            |            |            |
|               | ACTCCAGCCA |            |            |            |            |
| -             | ACCAGGCCAT |            |            |            |            |
|               | TCTCCCCTGA |            |            |            |            |
|               | TGAACACCAT |            |            |            |            |
|               | CCATCAATGA |            |            |            |            |
|               | CCCCGGCCA  |            |            |            |            |
| TCCACCCTCC    | AGGAGCAGAT | TGGCTGGATG | ACCAACAACC | CCCCCATCCC | TGTGGGGGAA |
| ATCTACAAGA    | GGTGGATCAT | CCTGGGCCTG | AACAAGATTG | TGAGGATGTA | CTCCCCCACC |
| TCCATCCTGG    | ACATCAGGCA | GGGCCCCAAG | GAGCCCTTCA | GGGACTATGT | GGACAGGTTC |
| TACAAGACCC    | TGAGGGCTGA | GCAGGCCTCC | CAGGAGGTGA | AGAACTGGAT | GACAGAGACC |
| CTGCTGGTGC    | AGAATGCCAA | CCCTGACTGC | AAGACCATCC | TGAAGGCCCT | GGGCCCTGCT |
| GCCACCCTGG    | AGGAGATGAT | GACAGCCTGC | CAGGGGGTGG | GGGGCCCTGG | TCACAAGGCC |
| AGGGTGCTGG    | CTGAGGCCAT | GTCCCAGGTG | ACCAACTCCG | CCACCATCAT | GATGCAGAGG |
| GGCAACTTCA    | GGAACCAGAG | GAAGACAGTG | AAGTGCTTCA | ACTGTGGCAA | GGTGGGCCAC |
| ATTGCCAAGA    | ACTGTAGGGC | CCCCAGGAAG | AAGGGCTGCT | GGAAGTGTGG | CAAGGAGGGC |
| CACCAGATGA    | AGGACTGCAA | TGAGAGGCAG | GCCAACTTCC | TGGGCAAAAT | CTGGCCCTCC |
| CACAAGGGCA    | GGCCTGGCAA | CTTCCTCCAG | TCCAGGCCTG | AGCCCACAGC | CCCTCCCGAG |
| GAGTCCTTCA    | GGTTTGGGGA | GGAGAAGACC | ACCCCCAGCC | AGAAGCAGGA | GCCCATTGAC |
| AAGGAGCTGT    | ACCCCCTGGC | CTCCCTGAGG | TCCCTGTTTG | GCAACGACCC | CTCCTCCCAG |
| ATGGCTCCCA    | TCTCCCCCAT | TGAGACTGTG | CCTGTGAAGC | TGAAGCCTGG | CATGGATGGC |
| CCCAAGGTGA    | AGCAGTGGCC | CCTGACTGAG | GAGAAGATCA | AGGCCCTGGT | GGAAATCTGC |
| ACTGAGATGG    | AGAAGGAGGG | CAAAATCTCC | AAGATTGGCC | CCGAGAACCC | CTACAACACC |
| CCTGTGTTTG    | CCATCAAGAA | GAAGGACTCC | ACCAAGTGGA | GGAAGCTGGT | GGACTTCAGG |
| GAGCTGAACA    | AGAGGACCCA | GGACTTCTGG | GAGGTGCAGC | TGGGCATCCC | CCACCCCGCT |
| GGCCTGAAGA    | AGAAGAAGTC | TGTGACTGTG | CTGGCTGTGG | GGGATGCCTA | CTTCTCTGTG |
| CCCCTGGATG    | AGGACTTCAG | GAAGTACACT | GCCTTCACCA | TCCCCTCCAT | CAACAATGAG |
|               |            |            |            |            | CTCCCCTGCC |
|               |            |            |            |            | CCCTGACATT |
|               |            |            |            |            | TGGGCAGCAC |
|               |            |            |            |            | CACCCCTGAC |
|               |            |            |            |            | CCCCGACAAG |
|               |            |            |            |            | TGACATCCAG |
|               |            |            |            |            | GGTGAGGCAG |
|               |            |            |            |            | GACTGAGGAG |
| GCTGAGCTGG    | AGCTGGCTGA | GAACAGGGAG | ATCCTGAAGG | AGCCTGTGCA | TGGGGTGTAC |

## FIGURE 33B

| TATO | ACCCCT  | CCAAGGACCT | GATTGCTGAG | ATCCAGAAGC | AGGGCCAGGG | CCAGTGGACC |
|------|---------|------------|------------|------------|------------|------------|
| TACC | TOTAAA  | ACCAGGAGCC | CTTCAAGAAC | CTGAAGACTG | GCAAGTATGC | CAGGATGAGG |
| GGGG | CCCACA  | CCAATGATGT | GAAGCAGCTG | ACTGAGGCTG | TGCAGAAGAT | CACCACTGAG |
| TCCA | TTGTGA  | TCTGGGGCAA | GACCCCCAAG | TTCAAGCTGC | CCATCCAGAA | GGAGACCTGG |
| GAG  | CCTGGT  | GGACTGAGTA | CTGGCAGGCC | ACCTGGATCC | CTGAGTGGGA | GTTTGTGAAC |
| ACCC | ccccc   | TGGTGAAGCT | GTGGTACCAG | CTGGAGAAGG | AGCCCATTGT | GGGGGCTGAG |
| ACC' | TCTATG  | TGGCTGGGGC | TGCCAACAGG | GAGACCAAGC | TGGGCAAGGC | TGGCTATGTG |
| ACC  | ACAGGG  | GCAGGCAGAA | GGTGGTGACC | CTGACTGACA | CCACCAACCA | GAAGACTGCC |
| CTC  | AGGCCA  | TCTACCTGGC | CCTCCAGGAC | TCTGGCCTGG | AGGTGAACAT | TGTGACTGCC |
| TCC  | CAGTATG | CCCTGGGCAT | CATCCAGGCC | CAGCCTGATC | AGTCTGAGTC | TGAGCTGGTG |
| AAC  | CAGATCA | TTGAGCAGCT | GATCAAGAAG | GAGAAGGTGT | ACCTGGCCTG | GGTGCCTGCC |
| CAC  | AAGGGCA | TTGGGGGCAA | TGAGCAGGTG | GACAAGCTGG | TGTCTGCTGG | CATCAGGAAG |
| GTG  | CTGTTCC | TGGATGGCAT | TGACAAGGCC | CAGGATGAGC | ATGAGAAGTA | CCACTCCAAC |
| TGG  | AGGGCTA | TGGCCTCTGA | CTTCAACCTG | CCCCCTGTGG | TGGCTAAGGA | GATTGTGGCC |
| TCC  | rgtgaca | AGTGCCAGCT | GAAGGGGGAG | GCCATGCATG | GGCAGGTGGA | CTGCTCCCCT |
| GGC  | ATCTGGC | AGCTGGCCTG | CACCCACCTG | GAGGGCAAGG | TGATCCTGGT | GGCTGTGCAT |
| GTG  | CCTCCG  | GCTACATTGA | GGCTGAGGTG | ATCCCTGCTG | AGACAGGCCA | GGAGACTGCC |
| TAC  | TTCCTGC | TGAAGCTGGC | TGGCAGGTGG | CCTGTGAAGA | CCATCCACAC | TGCCAATGGC |
| TCC  | AACTTCA | CTGGGGCCAC | AGTGAGGGCT | GCCTGCTGGT | GGGCTGGCAT | CAAGCAGGAG |
| TTT  | GGCATCC | CCTACAACCC | CCAGTCCCAG | GGGGTGGTGG | CCTCCATGAA | CAAGGAGCTG |
| AAG  | AAGATCA | TTGGGCAGGT | GAGGGACCAG | GCTGAGCACC | TGAAGACAGC | TGTGCAGATG |
| GCT  | GTGTTCA | TCCACAACTT | CAAGAGGAAG | GGGGGCATCG | GGGGCTACTC | CGCTGGGGAG |
| AGG. | attgtgg | ACATCATTGC | CACAGACATC | CAGACCAAGG | AGCTCCAGAA | GCAGATCACC |
| AAG  | ATCCAGA | ACTTCAGGGT | GTACTACAGG | GACTCCAGGA | ACCCCCTGTG | GAAGGGCCCT |
| GCC  | AAGCTGC | TGTGGAAGGG | GGAGGGGGCT | GTGGTGATCC | AGGACAACTC | TGACATCAAG |
| GTG  | GTGCCCA | GGAGGAAGGC | CAAGATCATC | AGGGACTATO | GCAAGCAGAT | GGCTGGGGAT |
| GAC  | TGTGTGG | CCTCCAGGCA | GGATGAGGAC | TAA        |            |            |
| SEQ  | ID NO:  | 38         |            |            |            |            |
|      |         |            |            |            |            |            |

#### FIGURE 34A

Met Gly Ala Arg Ala Ser Val Leu Ser Gly Gly Glu Leu Asp Lys Trp Glu Lys Ile Arg Leu Arg Pro Gly Gly Lys Lys Lys Tyr Lys Leu Lys His Ile Val Trp Ala Ser Arg Glu Leu Glu Arg Phe Ala Val Asn Pro Gly Leu Leu Glu Thr Ser Glu Gly Cys Arg Gln Ile Leu Gly Gln Leu Gln Pro Ser Leu Gln Thr Gly Ser Glu Glu Leu Arg Ser Leu Tyr Asn Thr Val Ala Thr Leu Tyr Cys Val His Gln Lys Ile Asp Val Lys Asp Thr Lys Glu Ala Leu Glu Lys Ile Glu Glu Glu Gln Asn Lys Ser Lys Lys Lys Ala Gln Gln Ala Ala Gly Thr Gly Asn Ser Ser Gln Val Ser Gln Asn Tyr Pro Ile Val Gln Asn Leu Gln Gly Gln Met Val His Gln Ala Ile Ser Pro Arg Thr Leu Asn Ala Trp Val Lys Val Val Glu Glu Lys Ala Phe Ser Pro Glu Val Ile Pro Met Phe Ser Ala Leu Ser Glu Gly Ala Thr Pro Gln Asp Leu Asn Thr Met Leu Asn Thr Val Gly Gly His Gln Ala Ala Met Gln Met Leu Lys Glu Thr Ile Asn Glu Glu Ala Ala Glu Trp Asp Arg Leu His Pro Val His Ala Gly Pro Ile Ala Pro Gly Gln Met Arg Glu Pro Arg Gly Ser Asp Ile Ala Gly Thr Thr Ser Thr Leu Gln Glu Gln Ile Gly Trp Met Thr Asn Asn Pro Pro Ile Pro Val Gly Glu Ile Tyr Lys Arg Trp Ile Ile Leu Gly Leu Asn Lys Ile Val Arg Met Tyr Ser Pro Thr Ser Ile Leu Asp Ile Arg Gln Gly Pro Lys Glu Pro Phe Arg Asp Tyr Val Asp Arg Phe Tyr Lys Thr Leu Arg Ala Glu Gln Ala Ser Gln Glu Val Lys Asn Trp Met Thr Glu Thr Leu Leu Val Gln Asn Ala Asn Pro Asp Cys Lys Thr Ile Leu Lys Ala Leu Gly Pro Ala Ala Thr Leu Glu Glu Met Met Thr Ala Cys Gln Gly Val Gly Gly Pro Gly His Lys Ala Arg Val Leu Ala Glu Ala Met Ser Gln Val Thr Asn Ser Ala Thr Ile Met Met Gln Arg Gly Asn Phe Arg Asn Gln Arg Lys Thr Val Lys Cys Phe Asn Cys Gly Lys Val Gly His Ile Ala Lys Asn Cys Arg Ala Pro Arg Lys Lys Gly Cys Trp Lys Cys Gly Lys Glu Gly His Gln Met Lys Asp Cys Asn Glu Arg Gln Ala Asn Phe Leu Gly Lys Ile Trp Pro Ser His Lys Gly Arg Pro Gly Asn Phe Leu Gln Ser Arg Pro Glu Pro Thr Ala Pro Pro Glu Glu Ser Phe Arg Phe Gly Glu Glu Lys Thr Thr Pro Ser Gln Lys Gln Glu Pro Ile Asp Lys Glu Leu Tyr Pro Leu Ala Ser Leu Arg Ser Leu Phe Gly Asn Asp Pro Ser Ser Gln Met Ala Pro Ile Ser Pro Ile Glu Thr Val Pro Val Lys Leu Lys Pro Gly Met Asp Gly Pro Lys Val Lys Gln Trp Pro Leu Thr Glu Glu Lys Ile Lys Ala Leu Val Glu Ile Cys Thr Glu Met Glu Lys Glu Gly Lys Ile Ser Lys Ile Gly Pro Glu Asn Pro Tvr Asn Thr Pro Val Phe Ala Ile Lys Lys Lys Asp Ser Thr Lys Trp Arg Lys Leu Val Asp Phe Arg Glu Leu Asn Lys Arg Thr Gln Asp Phe Trp Glu Val Gln Leu Gly Ile Pro His Pro Ala Gly Leu Lys Lys Lys Ser Val Thr Val Leu Ala Val Gly Asp Ala Tyr Phe Ser Val Pro Leu Asp Glu Asp Phe Arg Lys Tyr Thr Ala Phe Thr Ile Pro Ser Ile Asn Asn Glu Thr Pro Gly Ile Arg Tyr Gln Tyr Asn Val Leu Pro Gln Gly Trp Lys Gly Ser Pro Ala Ile Phe Gln Ser Ser Met Thr Lys Ile Leu Glu Pro Phe Arg Lys Gln Asn Pro Asp Ile Val Ile Tyr Gln Tyr Met Ala Ala Leu Tyr Val Gly Ser Asp Leu Glu Ile Gly Gln His Arg Thr Lys Ile Glu Glu Leu Arg Gln His Leu Leu Arg Trp Gly Leu Thr Thr Pro Asp Lys Lys His Gln Lys Glu Pro Pro Phe Leu Trp Met Gly Tyr Glu Leu His Pro

#### FIGURE 34B

Asp Lys Trp Thr Val Gln Pro Ile Val Leu Pro Glu Lys Asp Ser Trp Thr Val Asn Asp Ile Gln Lys Leu Val Gly Lys Leu Asn Trp Ala Ser Gln Ile Tyr Pro Gly Ile Lys Val Arg Gln Leu Cys Lys Leu Leu Arg Gly Thr Lys Ala Leu Thr Glu Val Ile Pro Leu Thr Glu Glu Ala Glu Leu Glu Leu Ala Glu Asn Arg Glu Ile Leu Lys Glu Pro Val His Gly Val Tyr Tyr Asp Pro Ser Lys Asp Leu Ile Ala Glu Ile Gln Lys Gln Gly Gln Gly Gln Trp Thr Tyr Gln Ile Tyr Gln Glu Pro Phe Lys Asn Leu Lys Thr Gly Lys Tyr Ala Arg Met Arg Gly Ala His Thr Asn Asp Val Lys Gln Leu Thr Glu Ala Val Gln Lys Ile Thr Thr Glu Ser Ile Val Ile Trp Gly Lys Thr Pro Lys Phe Lys Leu Pro Ile Gln Lys Glu Thr Trp Glu Thr Trp Trp Thr Glu Tyr Trp Gln Ala Thr Trp Ile Pro Glu Trp Glu Phe Val Asn Thr Pro Pro Leu Val Lys Leu Trp Tyr Gln Leu Glu Lys Glu Pro Ile Val Gly Ala Glu Thr Phe Tyr Val Ala Gly Ala Ala Asn Arg Glu Thr Lys Leu Gly Lys Ala Gly Tyr Val Thr Asn Arg Gly Arg Gln Lys Val Val Thr Leu Thr Asp Thr Thr Asn Gln Lys Thr Ala Leu Gln Ala Ile Tyr Leu Ala Leu Gln Asp Ser Gly Leu Glu Val Asn Ile Val Thr Ala Ser Gln Tyr Ala Leu Gly Ile Ile Gln Ala Gln Pro Asp Gln Ser Glu Ser Glu Leu Val Asn Gln Ile Ile Glu Gln Leu Ile Lys Lys Glu Lys Val Tyr Leu Ala Trp Val Pro Ala His Lys Gly Ile Gly Gly Asn Glu Gln Val Asp Lys Leu Val Ser Ala Gly Ile Arg Lys Val Leu Phe Leu Asp Gly Ile Asp Lys Ala Gln Asp Glu His Glu Lys Tyr His Ser Asn Trp Arg Ala Met Ala Ser Asp Phe Asn Leu Pro Pro Val Val Ala Lys Glu Ile Val Ala Ser Cys Asp Lys Cys Gln Leu Lys Gly Glu Ala Met His Gly Gln Val Asp Cys Ser Pro Gly Ile Trp Gln Leu Ala Cys Thr His Leu Glu Gly Lys Val Ile Leu Val Ala Val His Val Ala Ser Gly Tyr Ile Glu Ala Glu Val Ile Pro Ala Glu Thr Gly Gln Glu Thr Ala Tyr Phe Leu Leu Lys Leu Ala Gly Arg Trp Pro Val Lys Thr Ile His Thr Ala Asn Gly Ser Asn Phe Thr Gly Ala Thr Val Arg Ala Ala Cys Trp Trp Ala Gly Ile Lys Gln Glu Phe Gly Ile Pro Tyr Asn Pro Gln Ser Gln Gly Val Val Ala Ser Met Asn Lys Glu Leu Lys Lys Ile Ile Gly Gln Val Arg Asp Gln Ala Glu His Leu Lys Thr Ala Val Gln Met Ala Val Phe Ile His Asn Phe Lys Arg Lys Gly Gly Jle Gly Gly Tyr Ser Ala Gly Glu Arg Ile Val Asp Ile Ile Ala Thr Asp Ile Gln Thr Lys Glu Leu Gln Lys Gln Ile Thr Lys Ile Gln Asn Phe Arg Val Tyr Tyr Arg Asp Ser Arg Asn Pro Leu Trp Lys Gly Pro Ala Lys Leu Leu Trp Lys Gly Glu Gly Ala Val Val Ile Gln Asp Asn Ser Asp Ile Lys Val Val Pro Arg Lys Ala Lys Ile Ile Arg Asp Tyr Gly Lys Gln Met Ala Gly Asp Asp Cys Val Ala Ser Arg Gln Asp Glu Asp SEQ ID NO: 39

International application No.

|  |  | PC17US01/28                            | 501  |  |  |  |
|--|--|--|--|--|--|--|
|  | SSIFICATION OF SUBJECT MATTER  |  |  |  |  |  |
| IPC(7)   | : C12N 15/86<br>: 435/456  |  |  |  |  |  |
| US CL  | : 435/456 International Patent Classification (IPC) or to both n                         | etional classification and IPC         |  |  |  |  |
|  | DS SEARCHED  | adopai classification and it c         |  |  |  |  |
| Minimum do   | cumentation searched (classification system followed                                     | hy electification graphele)            |  |  |  |  |
|  | 24/205.1, 207.1, 227.1, 233.1; 435/69.1, 69.3, 173.                                      | •                                      |  |  |  |  |
| 0.0 12   |  | 2, 200.1, 020.1, 400, 000.20.72,       |  |  |  |  |
|  |  |  |  |  |  |  |
| Documentatio   | on searched other than minimum documentation to the                                      | e extent that such documents are inch  | ided in the fields searched  |  |  |  |
|  |  |  |  |  |  |  |
|  | · · · · · · · · · · · · · · · · · · ·  |  | <del></del>  |  |  |  |
|  | ta base consulted during the international search (nam                                   | ne of data base and, where practicable | e, search terms used)  |  |  |  |
| Please See Co  | ontinuation Sheet  |  |  |  |  |  |
|  |  |  |  |  |  |  |
| C. DOC   | UMENTS CONSIDERED TO BE RELEVANT   |  |  |  |  |  |
| Category *   | Citation of document, with indication, where ap  | propriate, of the relevant passages    | Relevant to claim No.  |  |  |  |
| х  | WO 96/39178 (ERTL et al.) 12 December 1996 (12   | .12.1996), see page 5, 6,10, 12, 13    | 1-3, 8-11, 18  |  |  |  |
|  | and claims 1 and 5.  |  |  |  |  |  |
| Y  |  |  | 4, 5, 13-17, 29-32, 34,<br>35, 37  |  |  |  |
| х  | US 6,019,978 A (ERTL et al.) 1 February 2000,(01   | /02/2000), see columns 2, 7 and 8.     | 1-3, 8-11, 18  |  |  |  |
|  | ,  | .,                                     |  |  |  |  |
| Y  |  |  | 4, 5, 13-17, 29-32, 34,<br>35, 37  |  |  |  |
| X,P  | US 6,287,571 6 (ERTL et al.) 11 September 200 and claim 1.                               | 1 (11/09/2001), see columns 2, 7, 8    | 1, 9, 18   |  |  |  |
| x  | US 5,643,579A (HUNG et al.) 1 July 1997 (01/07/  | 1997), see examples 1, 2, 25 and 26.   | 1-3, 8, 9-11, 18   |  |  |  |
| _  | -  |  | 4 5 12 17 20 22 24   |  |  |  |
| Y  |  |  | 4,5,13-17, 29-32, 34,<br>35, 37  |  |  |  |
| Y  | WANG et al. The use of an E1-deleted, replication  | -defective adenovirus recombinant      | 1-3, 9-11, 13-18   |  |  |  |
| -  | expressing the rabies virus glycoprotein for early vi                                    |  |  |  |  |  |
|  | Journal of Virology (March 1997) Vol. 71, No. 5,   | рр 3677-3683.                          |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
| Burther  | r documents are listed in the continuation of Box C.                                     | See patent family annex.               |  |  |  |  |
| <u> </u>   | pocial caregories of cited documents:  | "T"   limer document published after t | he international filing date or  |  |  |  |
|  |  | priority date and not in conflict      | with the application but cited to  |  |  |  |
|  | s defining the general state of the art which is not considered to<br>rticular relevance | understand the principle or theo       | ry underlying the invention  |  |  |  |
| _  |  |  | e; the claimed invention cannot be<br>considered to involve an inventive |  |  |  |
| date   | oplication or patent published on or after the international filing                      | step when the document is taken        |  |  |  |  |
| "L" document   | t which may throw doubts on priority claim(s) or which is cited                          | "Y" document of particular relevance   | e; the claimed invention cannot be                                       |  |  |  |
| to establi   | ish the publication date of another citation or other special reason                     | considered to involve an invent        | ive step when the document is  |  |  |  |
| (2s specified) combined with one or more other such documents, such combination being obvious to a person skilled in the art |  |  |  |  |  |  |
| O. quantum   | t referring to an oral disclosure, use, exhibition or other means                        | *&* document member of the same ;      | oatent family  |  |  |  |
|  | n published prior to the international filing date but later than the                    |  |  |  |  |  |
|  | actual completion of the international search  | Date of mailing of the international   | search report  |  |  |  |
| Of Ealman  | 2007 (05 02 2007)  | <b>19</b> AUG 200 <b>2</b> /           |  |  |  |  |
|  | 2002 (06.02.2002)<br>uailing address of the ISA/US                                       | Authorized officer                     | 1/10   |  |  |  |
|  | missioner of Patents and Trademarks  | 101                                    | lalkens to   |  |  |  |
|  | t PCT<br>shington, D.C. 20231  | Ulrike Winkler, Ph.D.                  |  |  |  |  |
|  | o. (703)305-3230   | Telephone No. 703-308-0196             | 1)   |  |  |  |
|  | · · · · · · · · · · · · · · · · · · ·  | <u> </u>                               |  |  |  |  |

Form PCT/ISA/210 (second sheet) (July 1998)

International application No.

PCT/US01/28861

| ategory * | Citation of document, with indication, where appropriate, of the relevant passages  | Relevant to claim No |
|-----------|---|----------------------|
| <b>Y</b>  | NATUK et al. Immunogenicity of recombinant human adenovirus -human immunodeficiency virus vaccines in chimpanzees. Aids Research and Human Retroviruses (1993) Vol. 9, No. 5, pp395-404, see material and methods.  | 1, 9, 29-32          |
| Y         | PREVEC et al. Immune response to HIV-1 gag antigens induced by recombinant adenovirus vectors in mice and rhesus macaque monkeys. Journal of Acquired Immune Deficincy Syndrome. (1991) Vol. 4, No. 6 pp. 568-76, see abstract.                           | 1, 9, 29-32          |
| Y         | LORI et al. Rapid protection against human immunodeficiency virus type 1 (HIV-1) replication mediated by high efficiency non-retroviral delivery of genes interfering with HIV-1 tat and gag. Gene Therapy (1994) Vol. 1, No. 1, pp. 27-31, see abstract. | 1,.9                 |
| Y         | PFARR et al. Differential effects of polyadenylation regions on gene expression in mammalian cells. DNA (1986) Vol. 5, No. 2, pp.115-22, see abstract.  | 16                   |
| Y         | NATUK et al. Adenovirus vectored vaccine. Developmental Biological Standards (1994)<br>Vol. 82, pp. 71-77, see abstract.  | 1,9                  |
|           |   |                      |
|           |   |                      |
|           | ·   |                      |
|           |   |                      |
|           |   |                      |
|           |   |                      |
|           |   |                      |
|           |   |                      |
|           |   |                      |
|           |   |                      |
|           |   |                      |
|           |   |                      |
|           |   |                      |

International application No.

PCT/US01/28861

| Box I Observations where certain claims were found unsearchable (Continuation of Item 1 of first sheet)   |
|---|
| This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:   |
| Claim Nos.: because they relate to subject matter not required to be searched by this Authority, namely:  |
| Claim Nos.:     because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:                           |
| 3. Claim Nos.:  because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).   |
| Box II Observations where unity of invention is lacking (Continuation of Item 2 of first sheet)   |
| This International Searching Authority found multiple inventions in this international application, as follows: Please See Continuation Sheet   |
| As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.  |
| 2. As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite  |
| payment of any additional fee.  3. As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:                 |
| 4. No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.: 1-5, 8-11, 13-18, 29-32, 34, 35, 37 |
| Remark on Protest  The additional search fees were accompanied by the applicant's protest.  No protest accompanied the payment of additional search fees.   |

International application No.

PCT/US01/28861

BOX II. OBSERVATIONS WHERE UNITY OF INVENTION IS LACKING

This application contains the following inventions or groups of inventions which are not so linked as to form a single general inventive concept under PCT Rule 13.1. In order for all inventions to be examined, the appropriate additional examination fees must be paid.

| Group | Claims  | ·  |
|-------|---|--|
| 1     | 1-5, 8-11,<br>13-18, 29,<br>30, 31, 32,<br>34, 35, 37 | The claims are directed to an adenoviral vector that is at least partially deleted of $\Delta E1$ , the vector contains the cis-acting packaging sequence of the wild type adenovirus genome, and a gene which encodes an HIV Gag protein (SEQ ID NO: 29) inserted in the parallel orientation of E1. In addition the vector contains a promoter and a polyadenylation signal. |
| 2     | 6, 7, 36  | The claims are directed to an adenoviral vector that is at least partially deleted of $\Delta E1$ and $\Delta E3$ , the vector contains the cis-acting packaging sequence of the wild type adenovirus genome, and a gene which encodes an HIV Gag protein (SEQ ID NO: 29).   |
| 3     | 12, 33  | The claims are directed to an adenoviral vector that is at least partially deleted of $\Delta E1$ , the vector contains the cis-acting packaging sequence of the wild type adenovirus genome, and a gene which encodes an HIV protein inserted in the antiparallel orientation of E1.  |
| 4     | 19-23, 38-42  | The claims are directed to a method of making and harvesting of a recombinant adenoviral particle that contains a gene encoding an HIV Gag protein.  |
| 5     | 24, 27, 28,<br>43, 46, 47                             | The claim is directed to a method of generating a cellular mediated immune response to HIV Gag protein with the recombinant adenoviral particle.   |
| 6     | 25, 26, 44,<br>45                                     | The claim is directed to a method of generating a cellular mediated immune response to HIV Gag protein with the recombinant adenoviral particle in addition to administering a DNA plasmid vaccine.  |
| 7     | 48-51, 53,<br>54, 56                                  | The claims are directed to an adenoviral vector that is at least partially deleted of $\Delta E1$ , the vector contains the cis-acting packaging sequence of the wild type adenovirus genome, and a gene which encodes an HIV Pol protein (SEQ ID NO: 1) inserted in the parallel orientation of E1.   |
| 8     | 48-51, 53,<br>54, 56                                  | The claims are directed to an adenoviral vector that is at least partially deleted of $\Delta E1$ , the vector contains the cis-acting packaging sequence of the wild type adenovirus genome, and a gene which encodes an HIV Pol protein (SEQ ID NO: 5) inserted in the parallel orientation of E1.   |
| 9     | 48-51, 53,<br>54, 56                                  | The claims are directed to an adenoviral vector that is at least partially deleted of<br><u>AE1</u> , the vector contains the cis-acting packaging sequence of the wild type adenovirus genome, and a gene which encodes an <u>HIV Pol protein (SEQ ID NO: 7)</u> inserted in the parallel orientation of E1.  |
| 10    | 52  | The claim is directed to an adenoviral vector that is at least partially deleted of $\Delta E1$ , the vector contains the cis-acting packaging sequence of the wild type adenovirus genome, and a gene which encodes an HIV Pol protein (SEQ ID NO: 1) inserted in the antiparallel orientation of E1.   |
| 11    | 52  | The claim is directed to an adenoviral vector that is at least partially deleted of $\Delta E1$ , the vector contains the cis-acting packaging sequence of the wild type adenovirus genome, and a gene which encodes an HIV Pol protein (SEQ ID NO: 5) inserted in the antiparallel orientation of E1.   |
| 12    | 52  | The claim is directed to an adenoviral vector that is at least partially deleted of $\Delta E1$ , the vector contains the cis-acting packaging sequence of the wild type adenovirus genome, and a gene which encodes an HIV Pol protein (SEQ ID NO: 7) inserted in the antiparallel orientation of E1.   |
| 13    | 55  | The claim is directed to an adenoviral vector that is at least partially deleted of $\Delta E1$  |

Form PCT/ISA/210 (second sheet) (July 1998)

International application No.

PCT/US01/28861

|     |                      | and ΔΕ3, the vector contains the cis-acting packaging sequence of the wild type adenovirus genome, and a gene which encodes an HIV Pol protein (SEQ ID NO: 1)   |
|-----|----------------------|---|
| 14  | 55                   | inserted in E1.  The claim is directed to an adenoviral vector that is at least partially deleted of $\Delta E1$  |
| • • |                      | and ΔE3, the vector contains the cis-acting packaging sequence of the wild type adenovirus genome, and a gene which encodes an HIV Pol protein (SEQ ID NO: 5) inserted in E1.   |
| 15  | .55                  | The claim is directed to an adenoviral vector that is at least partially deleted of $\Delta E1$ and $\Delta E3$ , the vector contains the cis-acting packaging sequence of the wild type adenovirus genome, and a gene which encodes an HIV Pol protein (SEQ ID NO: 7) inserted in E1.                                      |
| 16  | 57-61                | The claims are directed to a method of making and harvesting of a recombinant adenoviral particle that contains a gene encoding an HIV Pol protein.   |
| 17  | 62, 65, 66           | The claim is directed to a method of generating a cellular mediated immune response to HIV Pol protein with the recombinant adenoviral particle.  |
| 18  | 63, 64               | The claim is directed to a method of generating a cellular mediated immune response to HIV Pol protein with the recombinant adenoviral particle in addition to administering a DNA plasmid vaccine.   |
| 19  | 67-70, 72,<br>73, 75 | The claims are directed to an adenoviral vector that is at least partially deleted of<br><u>\Delta E1</u> , the vector contains the cis-acting packaging sequence of the wild type<br>adenovirus genome, and a gene which encodes an HIV Nef protein (SEQ ID NO: 9)<br>inserted in the parallel orientation of E1.          |
| 20  | 67-70, 72,<br>73, 75 | The claims are directed to an adenoviral vector that is at least partially deleted of<br>\$\Delta \text{E1}\$, the vector contains the cis-acting packaging sequence of the wild type adenovirus genome, and a gene which encodes an HIV Nef protein (SEQ ID NO: 11) inserted in the parallel orientation of \$\text{E1}\$. |
| 21  | 67-70, 72,<br>73, 75 | The claims are directed to an adenoviral vector that is at least partially deleted of<br><u>\Delta E1</u> , the vector contains the cis-acting packaging sequence of the wild type adenovirus genome, and a gene which encodes an <u>HIV Nef protein (SEQ ID NO: 13)</u> inserted in the parallel orientation of E1.        |
| 22  | 67-70, 72,<br>73, 75 | The claims are directed to an adenoviral vector that is at least partially deleted of<br><u>AE1</u> , the vector contains the cis-acting packaging sequence of the wild type<br>adenovirus genome, and a gene which encodes an <u>HIV Nef protein (SEQ ID NO: 15)</u><br>inserted in the parallel orientation of E1.        |
| 23  | 71                   | The claim is directed to an adenoviral vector that is at least partially deleted of $\Delta E1$ , the vector contains the cis-acting packaging sequence of the wild type adenovirus genome, and a gene which encodes an HIV Nef protein (SEQ ID NO: 9) inserted in the antiparallel orientation of E1.                      |
| 24  | 71                   | The claim is directed to an adenoviral vector that is at least partially deleted of $\Delta E1$ , the vector contains the cis-acting packaging sequence of the wild type adenovirus genome, and a gene which encodes an HIV Nef protein (SEQ ID NO: 11) inserted in the antiparallel orientation of E1.                     |
| 25  | 71                   | The claim is directed to an adenoviral vector that is at least partially deleted of $\Delta E1$ , the vector contains the cis-acting packaging sequence of the wild type adenovirus genome, and a gene which encodes an HIV Nef protein (SEQ ID NO: 13) inserted in the antiparallel orientation of E1.                     |
| 26  | 71                   | The claim is directed to an adenoviral vector that is at least partially deleted of $\Delta E1$ , the vector contains the cis-acting packaging sequence of the wild type adenovirus genome, and a gene which encodes an HIV Nef protein (SEQ ID NO: 15) inserted in the antiparallel orientation of E1.                     |
| 27  | 74                   | The claim is directed to an adenoviral vector that is at least partially deleted of $\Delta E1$ and $\Delta E3$ , the vector contains the cis-acting packaging sequence of the wild type adenovirus genome, and a gene which encodes an HIV Nef protein (SEQ ID NO: 9) inserted in E1.                                      |
| 28  | 74                   | The claim is directed to an adenoviral vector that is at least partially deleted of $\Delta E1$ and $\Delta E3$ , the vector contains the cis-acting packaging sequence of the wild type adenovirus genome, and a gene which encodes an HIV Nef protein (SEQ ID NO: 11) inserted in E1.                                     |
| 29  | 74                   | The claim is directed to an adenoviral vector that is at least partially deleted of $\Delta E1$ and $\Delta E3$ , the vector contains the cis-acting packaging sequence of the wild type  |

International application No.

PCT/US01/28861

|    |             | adenovirus genome, and a gene which encodes an HIV Nef protein (SEQ ID NO: 13) inserted in E1.  |
|----|-------------|---|
| 30 | 74          | The claim is directed to an adenoviral vector that is at least partially deleted of $\Delta E1$ and $\Delta E3$ , the vector contains the cis-acting packaging sequence of the wild type adenovirus genome, and a gene which encodes an HIV Nef protein (SEQ ID NO: 15) inserted in E1. |
| 31 | 76-80       | The claims are directed to a method of making and harvesting of a recombinant adenoviral particle that contains a gene encoding an HIV Nef protein.   |
| 32 | 81, 84, 85  | The claims are directed to a method of generating a cellular mediated immune response to HIV Nef with the recombinant adenoviral particle.  |
| 33 | 82, 83      | The claims are directed to a method of generating a cellular mediated immune response to HIV Nef with the recombinant adenoviral particle in addition to administering a DNA plasmid vaccine.   |
| 34 | 86a         | The claim is drawn to a multivalent vaccine wherein gag, pol and nef are expressed from three individual vectors.   |
| 35 | 86b, 88, 89 | The claims are drawn to a multivalent vaccine wherein gag, pol and nef are expressed from one individual vectors.   |
| 36 | 86c, 88     | The claims are drawn to a multivalent vaccine wherein gag, pol and nef are expressed from two individual vectors, one expressing nef-pol fusion and one expressing gag.   |
| 37 | 86d, 87, 88 | The claims are drawn to a multivalent vaccine wherein gag, pol and nef are expressed from two individual vectors, one expressing gag-pol fusion and one expressing nef.   |
| 38 | 86e, 88     | The claims are drawn to a multivalent vaccine wherein gag, pol and nef are expressed from two individual vectors, one expressing nef-gag fusion and one expressing pol.   |
| 39 | 86f, 88     | The claims are drawn to a multivalent vaccine wherein gag, pol and nef are expressed from a single vectors as a fusion protein.   |
| 40 | 86g, 88     | The claims are drawn to a multivalent vaccine wherein gag and pol are expressed from two individual vectors.  |
| 41 | 86h, 88, 89 | The claims are drawn to a multivalent vaccine wherein gag and pol are expressed individually from one vector.   |
| 42 | 86i, 88     | The claims are drawn to a multivalent vaccine wherein pol and nef are expressed from two individual vectors.  |
| 43 | 86j, 88, 89 | The claims are drawn to a multivalent vaccine wherein pol and nef are expressed from individually from one vector.  |
| 44 | 86k, 88     | The claims are drawn to a multivalent vaccine wherein nef and gag are expressed individually from one vector.   |
| 45 | 861, 88, 89 | The claims are drawn to a multivalent vaccine wherein nef and gag are expressed individually from one vector.   |
| 46 | 86m, 88     | The claims are drawn to a multivalent vaccine wherein gag and pol are expressed as a fusion protein from one vector.  |
| 47 | 86n, 88     | The claims are drawn to a multivalent vaccine wherein pol and nef are expressed as a fusion protein from one vector.  |
| 48 | 86o, 88     | The claims are drawn to a multivalent vaccine wherein nef and gag are expressed as a  |

The inventions listed as Groups 1-48 do not relate to a single general inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons:

The technical feature linking groups 1-33 appears to be a recombinant adenoviral vector wherein the adenoviral vector is at least partially deleted in E1 but the vector may contain more deletions, the vector contains wild type sequences including packaging signals and a gene encoding a heterologous HIV protein or fragments thereof. Ertl et al. (WO 96/39178) disclose a recombinant adenoviral vector that is deleted in E1 and partially deleted in E3, the remainder of the adenoviral vector contains wild type sequences. The vector additionally contains an insertion of a heterologous protein which includes HIV proteins (see abstract and claims 1 and 5). Therefore, the technical feature linking the inventions of groups 1-45 does not constitute a special technical feature as defined by PCT Rule 13.2, as it does not define a contribution over the prior art.

The special technical feature of the following groups 1-3, 7-15, 19-30 and 34-48 is considered to be the combination of sequences that is disclosed in each group, see individual claim groupings above for the different sequences. The DNA disclosed in each group is made up of a different sequence having a different structure and different function.

International application No.

PCT/US01/28861

The special technical feature of group 4, 16 and 31 is considered to be a method of producing recombinant adenoviral particles. Each group contains different sequences hence the resulting particles would have different structures and functions associated with the particle.

The special technical feature of group 5, 17 and 32 is considered to be a method of producing a cellular mediated immune response to the heterologous protein encoded by the different adenoviral vectors. Each group contains different sequences a encoding different protein, therefore the resulting immune response will also be different.

The special technical feature of group 6, 18 and 33 is considered to be a method of producing a cellular mediated immune response to the heterologous protein encoded by the different adenoviral vectors in conjunction with immunizing the individual a DNA plasmid vaccine. Each method contains different sequences encoding a different protein, therefore the resulting immune response will also be different.

Accordingly, groups 1-48 are not so linked by the same or corresponding technical feature as to form a single general inventive concept.

# Continuation of B. FIELDS SEARCHED Item 3:

WEST 2.0, STN-BIOSIS, MEDLINE

adenoviral vector, deletion, HIV, Gag, polyadenylation signal, CMV promoter